

PREPARING FOR THE 2020 CENSUS: WILL THE TECHNOLOGY BE READY?

JOINT HEARING
BEFORE THE
SUBCOMMITTEE ON
GOVERNMENT OPERATIONS
AND THE
SUBCOMMITTEE ON
INFORMATION TECHNOLOGY
OF THE
COMMITTEE ON OVERSIGHT
AND GOVERNMENT REFORM
HOUSE OF REPRESENTATIVES
ONE HUNDRED FOURTEENTH CONGRESS
FIRST SESSION

NOVEMBER 3, 2015

Serial No. 114-82

Printed for the use of the Committee on Oversight and Government Reform



Available via the World Wide Web: <http://www.fdsys.gov>
<http://www.house.gov/reform>

U.S. GOVERNMENT PUBLISHING OFFICE

22-362 PDF

WASHINGTON : 2017

For sale by the Superintendent of Documents, U.S. Government Publishing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512-1800; DC area (202) 512-1800
Fax: (202) 512-2104 Mail: Stop IDCC, Washington, DC 20402-0001

COMMITTEE ON OVERSIGHT AND GOVERNMENT REFORM

JASON CHAFFETZ, Utah, *Chairman*

JOHN L. MICA, Florida
MICHAEL R. TURNER, Ohio
JOHN J. DUNCAN, JR., Tennessee
JIM JORDAN, Ohio
TIM WALBERG, Michigan
JUSTIN AMASH, Michigan
PAUL A. GOSAR, Arizona
SCOTT DESJARLAIS, Tennessee
TREY GOWDY, South Carolina
BLAKE FARENTHOLD, Texas
CYNTHIA M. LUMMIS, Wyoming
THOMAS MASSIE, Kentucky
MARK MEADOWS, North Carolina
RON DESANTIS, Florida
MICK MULVANEY, South Carolina
KEN BUCK, Colorado
MARK WALKER, North Carolina
ROD BLUM, Iowa
JODY B. HICE, Georgia
STEVE RUSSELL, Oklahoma
EARL L. "BUDDY" CARTER, Georgia
GLENN GROTHMAN, Wisconsin
WILL HURD, Texas
GARY J. PALMER, Alabama

ELIJAH E. CUMMINGS, Maryland, *Ranking
Minority Member*
CAROLYN B. MALONEY, New York
ELEANOR HOLMES NORTON, District of
Columbia
WM. LACY CLAY, Missouri
STEPHEN F. LYNCH, Massachusetts
JIM COOPER, Tennessee
GERALD E. CONNOLLY, Virginia
MATT CARTWRIGHT, Pennsylvania
TAMMY DUCKWORTH, Illinois
ROBIN L. KELLY, Illinois
BRENDA L. LAWRENCE, Michigan
TED LIEU, California
BONNIE WATSON COLEMAN, New Jersey
STACEY E. PLASKETT, Virgin Islands
MARK DeSAULNIER, California
BRENDAN F. BOYLE, Pennsylvania
PETER WELCH, Vermont
MICHELLE LUJAN GRISHAM, New Mexico

SEAN McLAUGHLIN, *Staff Director*

DAVID RAPALLO, *Minority Staff Director*

JEFF POST, *Government Operations Subcommittee Deputy Staff Director*

SHARON CASEY, *Deputy Chief Clerk*

SUBCOMMITTEE ON GOVERNMENT OPERATIONS

MARK MEADOWS, North Carolina, *Chairman*

JIM JORDAN, Ohio	GERALD E. CONNOLLY, Virginia, Ranking
TIM WALBERG, Michigan, <i>Vice Chair</i>	Minority Member
TREY GOWDY, South Carolina	CAROLYN B. MALONEY, New York
THOMAS MASSIE, Kentucky	ELEANOR HOLMES NORTON, District of
MICK MULVANEY, South Carolina	Columbia
KEN BUCK, Colorado	WM. LACY CLAY, Missouri
EARL L. "BUDDY" CARTER, Georgia	STACEY E. PLASKETT, Virgin Islands
GLENN GROTHMAN, Wisconsin	STEPHEN F. LYNCH, Massachusetts

SUBCOMMITTEE ON INFORMATION TECHNOLOGY

WILL HURD, Texas, *Chairman*

BLAKE FARENTHOLD, Texas, <i>Vice Chair</i>	ROBIN L. KELLY, Illinois, <i>Ranking Member</i>
MARK WALKER, North Carolina	GERALD E. CONNOLLY, Virginia
ROD BLUM, Iowa	TAMMY DUCKWORTH, Illinois
PAUL A. GOSAR, Arizona	TED LIEU, California

CONTENTS

Hearing held on November 3, 2015	Page 1
WITNESSES	
The Hon. John H. Thompson, Director, U.S. Census Bureau	
Oral Statement	6
Written Statement	9
Mr. Steven I. Cooper, Chief Information Officer, U.S. Department of Commerce	
Oral Statement	14
Written Statement	16
Mr. Robert Goldenkoff, Director, Strategic Issues, U.S. Government Accountability Office	
Oral Statement	23
Written Statement	25
Ms. Carol R. Cha, Director, Information Technology Acquisition Management Issues, U.S. Government Accountability Office	
Oral Statement	39
Written Statement	41
APPENDIX	
Rep. Plaskett Statement for Record	78
RESPONSE Director Thompson, Census, Questions for the Record	79
RESPONSE Mr. Cooper, CIO, DOC, Questions for the Record	100
RESPONSE Mr. Goldenkoff, GAO, Questions for the Record	108
RESPONSE Ms. Cha, GAO, Questions for the Record	113

PREPARING FOR THE 2020 CENSUS: WILL THE TECHNOLOGY BE READY?

Wednesday, November 3, 2015

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON GOVERNMENT OPERATIONS, JOINT
WITH THE SUBCOMMITTEE ON INFORMATION TECHNOLOGY,
COMMITTEE ON OVERSIGHT AND GOVERNMENT REFORM,
Washington, D.C.

The subcommittees met, pursuant to call, at 2:36 p.m., in Room 2154, Rayburn House Office Building, Hon. Mark Meadows [chairman of the Subcommittee on Government Operations] presiding.

Present from Subcommittee on Government Operations: Representatives Meadows, Walberg, Carter, Grothman, Connolly, Maloney, and Clay.

Present from Subcommittee on Information Technology: Representatives Hurd, Walker, Blum, Kelly, and Duckworth.

Mr. MEADOWS. The Subcommittee on Government Operations and the Subcommittee on Information Technology will come to order. And without objection, the chair is authorized to declare a recess at any time. Thank you so much for being here. My apologies for being a little late coming from votes.

We're here today to talk about a very important topic, but one that does not get much attention in the years that don't end in zero. And so as we talk about the Census, actually last month the Bureau released its operational plan for the 2020 Census. The purpose of this plan was to ensure that the Bureau has all the necessary systems and procedures in place to effectively complete its core mission.

As our witnesses' testimony will indicate, the plan for 2020 is probably the most ambitious Census plan in our Nation's history. Quite simply, the Bureau is attempting to rewrite how the Census is conducted from the ground up, making it easier for people to respond and better value to the American taxpayer certainly as we look at that.

These are laudable and necessary goals, particularly in light of the Bureau's track record. The 2010 Census cost some \$12.3 billion or more than \$100 per household. This is substantially more than the \$8 billion or the about \$70 per household cost of the 2000 year Census. In fact, we have to go all the way back to the 1970 Census to find a Census where the cost did not rise dramatically compared to its predecessor.

To correct this unsustainable trend, the Census Bureau is attempting to greatly increase its ability to leverage technology to utilize preexisting data in a completely new way. Assuming the

current design hold, this will be the first Census where Americans will be able to respond through the Internet. It will be the first where Nonresponse Followup enumerators will be able to collect data electronically and have the case list updated in real time and organize their route based on travel efficiency with directions. What a novel idea.

It will be the first time that the Bureau will use both Federal, State, and local addresses updates for the past decade to avoid the added expense of a nationwide, in-field address canvass. It will also be the first time that the Bureau will use the Federal administrative records like tax returns and Medicare information to identify vacant housing units and to encourage self-response. This plan involves sending paper reply forms only to the older individuals that are in areas with poor Internet connectivity.

If the Bureau can deliver on its plan, it's believed that the 2020 Census could cost less per household than the 2010 Census and almost the same on an absolute level. However, "if" is the key word in that last sentence, and in preparation for the 2010 Census, the Bureau spent some \$600 million attempting to build a proprietary handheld device for the enumerators to electronically report the field data. Despite this massive investment of hard-working American taxpayer dollars, poor program management led the Bureau to abandon that project in its entirety in 2008, forcing them to go back to paper response forms for the 2010 Census.

The Bureau tells us that it has learned valuable lessons from that failed effort. The initiatives it is hoping to implement for the 2020 Census are several times the size and complexity of the program that failed in 2008 and it is why we are here today. Although I know that there are still some questions about how exactly the Bureau will use the administrative records, I feel confident in saying that all of the members on both the subcommittees are broadly supportive of what the Bureau is trying to accomplish.

With that in mind, my goal today is to learn more about the Bureau's ongoing initiatives, but also make sure realistic time lines, deadlines, operational controls are in place to avoid the failures of the past. Given the current budget environment, we cannot afford another incident like happened in 2008.

But in closing, I would like to thank all of our witnesses. And some of you I've met before. And I'm encouraged to have you with us for your time today and I look forward to working with you as the clock ticks towards a Census date of April 1, 2020.

The chair now recognizes my good friend, the ranking member of the Subcommittee on Government Operations, Mr. Connolly, for his opening statement.

Mr. CONNOLLY. Thank you, Mr. Chairman. And thanks for this joint hearing, which is well-timed given the recent release of the 2020 operational plan for the Census.

At the midpoint of this decade it's instructive to look at lessons learned from the last Census and the preparations underway for the next. The data derived from the Census provides a treasure trove of information both to the public and private sectors.

For the government, Census data provides the foundation for our representative system of governance. From Congress down to local school boards, Census data is used to prudently and fairly allocate

Federal, State, and local dollars to the communities we represent. For businesses, this information informs employment and capital investment decisions, critical for the private sector.

The Census is one of the few government functions expressly mandated by the Constitution. By the way, so is the Postal Service. And the advent of technology holds the promise of making this once-in-a-decade enumeration less arduous and hopefully less costly.

Success will be largely dependent on the management and deployment of technology. The Bureau's botched attempt to deploy mobile handheld devices for the 2010 Census actually increased costs—increased costs—by \$3 billion. The 2010 Census was 56 percent more costly than its predecessor in 2000.

Today the Bureau is proposing to expand its use of technology, including new handheld devices, with the stated goal of lowering its latest estimated cost by nearly one-third. That is a noble goal. One of the new initiatives the Bureau believes will produce great efficiencies and savings is the commonsense step of cross-referencing data from other Federal and non-Federal sources, such as the IRS, the U.S. Postal Service, and others. A county-level test earlier this year showed using such records reduced the follow-up workload at homes known to be vacant by nearly 12 percent.

In addition, the Bureau intends to expand use of modern technology. I hope we hear today what lessons from the 2010 experience with handheld devices are being applied to this new effort.

I'm also pleased to finally welcome the Census to the 21st century by offering an online response option. In today's high-tech mobile society, people simply expect to be able to conduct business online and in many cases in the palm of their hand with a smart device.

I'm encouraged by Mr. Cooper's prepared remarks in which he indicates the Bureau is on track to meet its goals. As the Commerce Department's CIO, he is accountable for overseeing the Census Enterprise Data Collection and Processing initiative, which will integrate and standardize systems and data sharing across the Bureau.

I want to congratulate him on his efforts to implement the FITARA act—better known as Connolly-Issa, Issa-Connolly—which is guiding IT investments with enhanced risk assessments and performance metrics. Commerce has been a leader among Federal agencies for embracing the new IT management model, and I would welcome his insight on how it has helped—or not—improve preparations for the Census.

I also look forward to hearing responses from Mr. Cooper and Mr. Thompson, Director of the Census Bureau, to concerns raised by the GAO that the Bureau is not moving quickly enough to achieve certain milestones. For example, GAO has cited gaps in staffing and the deferral of key IT decisions, such as the scope of the IT infrastructure that will be necessary, cybersecurity protocols, and a procurement strategy for the handheld devices. These are important issues that must be resolved soon as the window of time in which the Bureau has to complete this transformation before end-to-end testing of the system is scheduled to begin in 2018.

As Mr. Cooper notes, however, the Bureau has delivered its operational plan 3 years earlier than it did in 2010. And, hopefully, that allows us time to identify and address gaps or shortfalls well in advance of the deadline.

Finally, Mr. Chairman, let me address the elephant in the room, which is the urgency for Congress adequately to fund the Census. Providing robust oversight of its operations and these IT investments is a laudable and necessary goal for us, but it is for naught if the Bureau does not have the resources it needs with which to execute its constitutional duty in what has now become a monumental task given the size of our population and the complexity of our country.

The initial budget proposal offered by my friends for fiscal year 2016 would have reduced funding for the Census Bureau by \$374 million, almost one-third compared to the President's request. I'm hopeful that last week's bipartisan budget agreement will allow us to restore those necessary investments. We cannot afford to short-change an activity so fundamental to our democracy and the sustained well-being of our communities.

I look forward to hearing from the panel. Thank you, Mr. Chairman.

Mr. MEADOWS. Thank you, Mr. Connolly.

The chair recognizes the chairman of the Information Technology Subcommittee, the gentleman from Texas, Mr. Hurd.

Mr. HURD. Thank you, Chairman. I would like to thank the distinguished gentlemen from North Carolina and Virginia for holding this important hearing today.

Every 10 years the United States Census Bureau administers a very important survey of the American public. It's vital for many reasons. We will use the gathered information to ensure the proper distribution of government funds and the proportionate number of representatives in government. But in order to make sure the information collected is as accurate as possible, the Census Bureau must deploy the best possible tactics and plans to collect the data.

I want to give credit where credit is due. The Census Bureau is attempting to go to great lengths to drive down their costs and the results will be significant. This is atypical of the culture in Washington where agencies always seem to be asking for more tax dollars without producing the results that warrant them. The Census Bureau has accomplished this by utilizing existing information and new technology, something that other Federal agencies could learn from.

But like many Federal agencies, the Census Bureau has serious IT challenges that must be addressed and corrected now. First, while the Bureau has made some significant progress in fully staffing its IT department, a number of key leadership positions remain vacant. These positions need to be filled immediately, and I'm looking forward to this hearing today on how the agency plans to address this.

Second, the Census Bureau hasn't addressed all the GAO recommendations on improving its IT systems. We saw what happens when agencies ignore IG reports with the massive data breach at OPM. This is a trend that I continually see as chairman of the Subcommittee on Information Technology, Federal agencies ignoring

IG and GAO recommendations. It is vital that the Census Bureau take action to address all of the remaining recommendations as soon as possible.

Finally, I agree wholeheartedly with the GAO and our witness, Ms. Cha, that the Bureau must begin now to make critical IT decisions. Their deferral of decisions in IT so far has, in my view, increased the overall risk around the 2020 Census. I look forward to hearing from all our witnesses today on specific ways we can successfully utilize technology to complete an accurate and cost-effective Census in 2020. I yield back.

Mr. MEADOWS. I thank the gentleman.

The chair now recognizes Ms. Kelly, the ranking member of the Subcommittee on Information Technology, for her opening statement.

Ms. KELLY. Thank you, Mr. Chair.

Good afternoon. Mr. Chairman, thank you for holding this important hearing. I would also like to extend my thanks to the witnesses here today.

The purpose of this oversight hearing is to determine if the Census Bureau's plans for 2020 Census is sufficient to ensure complete and accurate counts of the Nation's population. Getting an accurate count is important. Numerous key decisions are based on data collected from the decennial Census. Census data is integral in determining the equitable distribution of Federal funds, enforcing civil and voting rights legislation, and determining congressional proportionment, among other things.

With advances in technology, we have tools available to help administer the Census more efficiently and accurately. For instance, in October 2014, the Bureau began an enterprise-wide IT initiative called the Census Enterprise Data Collection and Processing Program, which is intended to, "deliver a system-of-systems to serve all of the Bureau's survey, data collection, and processing functions—rather than continuing to build and maintain unique survey-specific systems with redundant capabilities."

CEDCaP will help reengineer fieldwork by implementing an operational control system to track and manage field assignments. This will help census takers decide, using real-time data, which houses to visit on a daily basis.

CEDCaP will also be responsible for the development of a Web-based survey application in order to maximize Internet self-response. The 2020 Census will be the first Census with the option of responding online. The Bureau will rely heavily on Internet responses as part of its cost-saving initiatives.

The Bureau also decided to use mobile devices for field data collection purposes. The Bureau estimates it will save nearly \$400 million through the increased use of technology, which is certainly to be commended.

While technology can help increase efficiency, which yields significant savings, I am also concerned about the safety of the public's personal information. As you may know, the Bureau experienced a data breach in July. While no sensitive information was stolen, this incident underscores the importance of having controls in place to protect sensitive information.

I applaud the Bureau for responding swiftly to the breach and recognizing the need for continued monitoring of their systems. It is our job to ensure that these agencies have the resources they need to keep our private information safe from hackers.

Additionally, I share GAO's concern that a number of important IT decisions have yet to be made for the 2020 Census. GAO recommends that key IT decisions be made quickly or we'll find ourselves in the situation we saw in 2010, where a costly IT project went awry, causing turmoil for the entire Census operation.

I look forward to hearing from the Bureau regarding its ability to meet critical milestones as they finalize plans for 2020, and I look forward to a productive discussion on this vital issue.

I yield back the balance of my time, and thank you.

Mr. MEADOWS. Thank you, Ms. Kelly.

I will hold the record open for 5 legislative days for any member who would like to submit a written statement.

Mr. MEADOWS. We will now recognize our panel of witnesses. I'm pleased to welcome the Honorable John Thompson, the Director of the U.S. Census Bureau; Mr. Steven Cooper, Chief Information Officer at the U.S. Department of Commerce; Mr. Robert Goldenkoff, Director of Strategic Issues at the U.S. Government Accountability Office; and Ms. Carol Cha, Director of Information Technology Acquisitions Management Issues at the U.S. Government Accountability Office.

Welcome to you all.

Pursuant to committee rules, all witnesses will be sworn in before they testify. So if you would please rise and raise your right hand.

Do you solemnly swear or affirm that the testimony you're about to give will be the truth, the whole truth, and nothing but the truth? Thank you. Please be seated.

Let the record reflect that the witnesses answered in the affirmative.

And in order to allow time for discussion, we would appreciate if you would limit your oral testimony to 5 minutes. However, your entire written statement will be made part of the record.

Mr. Thompson, who I would also say looks an awful lot like Robert De Niro, you are recognized for 5 minutes.

WITNESS STATEMENTS

STATEMENT OF JOHN H. THOMPSON

Mr. THOMPSON. Thank you, Chairman Meadows.

Chairman Meadows, Chairman Hurd, Ranking Members Connolly and Kelly, thank you for the opportunity to testify this afternoon. I'm also pleased to be testifying with Mr. Steve Cooper, the Chief Information Officer of the Department of Commerce, with whom we work very closely.

I am honored to serve as the Director of the Census Bureau which has the tremendous responsibility of administering an accurate Census that fairly represents everyone in America. We have reached a major milestone in our preparations for the 2020 Census. Just last month, we released an operational plan that allows us to

change the way we have conducted the decennial Census for over 46 years and save \$5.2 billion.

For each Census since 1970, the paper-based process has been the standard, and it has been increasingly challenged by the growing diversity and complexity of our Nation. We do not believe that a paper-and-pencil approach to the Census is sustainable for the 2020 or future Censuses. The Census Bureau has concentrated on four innovation areas that have the potential to improve Census operations and save taxpayers money. I will briefly describe each of these innovation areas.

First, we're building a more accurate address list, and for 2020, we have reengineered the process to incorporate a 100 percent in-office canvassing that will be supplemented with up to a 25 percent in-field operation. We have already started this effort, and we expect to realize \$900 million in cost savings from this strategy.

Second, we want to make it easier for people to respond, and our strategy incorporates not only the Internet as the primary response option, but also an integrated communication and partnership strategy and tailored contact strategies. We will mail an invitation to all housing units encouraging Internet response.

For 28 percent of the addresses, however, in areas with low Internet connectivity, we will be including a paper questionnaire in the first mailing package. We will also incorporate a Census Questionnaire Assistance program with both telephone and Web-chat response options. We believe that by opening up the process and making it easy, we will encourage participation with respondents who have typically not responded. These efforts offer the potential of approximately \$400 million in savings.

Third, using existing information already provided to the government, we can reduce door-to-door visits in what we call the Non-response Followup operation. To reduce this workload, we are proposing to use administrative records for enumeration in two ways. We are planning to use administrative records to remove vacant addresses before sending census takers into the field, and we are also planning to use administrative records to enumerate occupied housing units after we have knocked on each door at least once. Using administrative records in this way will result in savings of \$1.4 billion.

Fourth, we are automating our field operations to save an estimated \$2.5 billion. We will now use mobile technology to achieve significant efficiencies in the 2020 Census. I should note that we have successfully developed an innovative prototype system. We are now able to provide optimized assignments to our interviewers, including daily route assignments, the best time of day to attempt contact. We are also able to provide the supervisors of these enumerators real-time updates and alerts regarding the progress of the workers they oversee.

Supporting these efforts is our enterprise approach to survey and Census data collection and processing through shared services, which we call CEDCaP. In the past, duplicative systems were created and used for every survey and Census. We will now move to a small suite of shared, reusable systems. Based on my experience in overseeing the 2000 Census and in the private sector, I am con-

fidant we are on course and we have a schedule for all major decisions.

Finally, we are committed to protecting the privacy and confidentiality of individual information. The Census Bureau has implemented a robust, comprehensive, and layered cybersecurity program that is constantly evaluated by experts.

The last 5 years have provided the groundwork for the 21st century Census. And now, looking forward, we must turn our attention to counting every person in America. But still a significant risk we face is receiving adequate funding. If adequate funding is not received in fiscal year 2016, we will prioritize activities to ensure that the 2018 end-to-end test will take place on time. We are committed to ensuring an accurate Census that fairly represents all people in America. If we have to defer activities to later years, the cost of the Census will increase.

I am confident the Census Bureau can achieve these objectives given congressional support, and I look forward to discussing the 2020 Census operational plan and other aspects of our planning with you today. Thank you.

[Prepared statement of Mr. Thompson follows:]



**Written Testimony
of
John H. Thompson
Director of the U.S. Census Bureau**

**Before the U.S. House of Representatives
Subcommittee on Government Operations
and the
Subcommittee on Information Technology
of the Committee on Government Oversight and Reform**

November 3, 2015

Chairman Meadows, Chairman Hurd and Ranking Members Connolly and Kelly, thank you for the opportunity to update the House Subcommittees on Government Operations and Information Technology on the U.S. Census Bureau's progress in planning for the 2020 Census. I am also pleased to be testifying with Mr. Steve Cooper, the Chief Information Officer of the Department of Commerce with whom we work closely.

I am honored to serve as the Director of the Census Bureau, which has the Constitutional responsibility of administering an accurate census that fairly represents every person living in America, and is fundamental to our democracy. We have reached a major milestone in our preparations for the 2020 Census. We have released an operational plan that allows us to change fundamentally the way in which the Census Bureau has conducted the Decennial Census for over 40 years. The 1970 Census was a breakthrough for its time. We built an address list and mailed questionnaires to each housing unit on the list. We asked respondents to complete and return the questionnaires through the mail. We developed automated processes to capture the information on the returns. However, the task of collecting information from those households that did not self-respond required recruiting and managing a army of enumerators using paper and pencil. For each census since 1970, this paper-based process has been the standard, and it has been increasingly challenged by the growing diversity and complexity of our nation. We do not believe that a paper and pencil approach to the Census is sustainable for the 2020 or future censuses. It is no longer affordable and more importantly, is not adequate to meet the challenges of enumerating an increasingly diverse society. The operational plan that we prepared incorporates a number of innovations that will allow us to achieve a modern census for the 21st century. We are proposing to use mobile technology, administrative records, innovations from the geospatial industry, and self-response via the Internet to build a census that will cost far less than repeating the outdated processes used in 2010 to accurately count every person living in America.

I will briefly discuss the components of our 2020 Census Operational Plan to provide an overview of the key operations and innovations it includes. We are also pleased that this plan has been released three years earlier than the corresponding plan for the 2010 Census. It includes a number of important decisions that are based on the important research we have conducted and clearly documents the process by which we will conduct additional development and testing to have all key decisions in place to support a complete end-to-end test of our methodologies in 2018.

The Census Bureau has concentrated on four key innovation areas: reengineering address canvassing; optimizing self-response; utilizing administrative records and third-party data; and reengineering field operations. Each of these combined innovation areas has the potential to dramatically improve census operations and to save taxpayers money. At this time, the 2020 Census is estimated to cost \$12.5 billion compared with a cost of \$17.8 billion that it would take to repeat the paper and pencil design of the 2010 Census. This represents approximately \$5.2 billion in cost avoidance.

I will briefly describe each of these innovation areas.

First, we must make sure we have a complete address list, and to accomplish this, we will canvass the entire nation as we have in each previous census. However, for 2020 we have reengineered the process to incorporate a one hundred percent in-office canvassing that will be supplemented with a twenty-five percent field operation. We have already started this effort and we expect to realize \$900 million in cost savings from this strategy.

Second, to generate the largest possible self-response, our strategy incorporates not only the Internet as the principal response option, but also a nationwide integrated communication and partnership strategy and tailored contact strategies. We will mail an invitation to all housing units, encouraging Internet response. For 20 percent of the addresses in areas with low Internet connectivity and older populations, we will be including a questionnaire in the first mail package. We will also incorporate a Census Questionnaire Assistance program with both telephone and web-chat options. We will make it easy for people to respond anytime and anywhere, which means we will accept responses over the Internet, even when respondents cannot locate and provide their unique census identification code. We believe that by opening up the process and making it easy, we will encourage participation with respondents who have typically not responded. These efforts offer the potential of approximately \$400 million in savings.

Third, using administrative records and third-party data will also increase our efficiency by reducing the workload that we face in collecting information from those housing units that do not provide a self-response – an operation we refer to as nonresponse follow-up. Our current estimate for the non-response follow-up workload for the 2020 Census is 56 million housing units. To reduce the workload, we are proposing to use administrative records for enumeration in two ways.

We are planning to use administrative records to remove the vacant addresses before sending census takers into the field, meaning enumerators will not knock on doors where no one is living, which represents a significant number of addresses, approximately 11 percent of the workload or six million housing units from the total workload of 56 million. We are also planning to use administrative records to enumerate occupied housing units after attempting to obtain a response through an in person visit. We will knock on every door at least once, which we anticipate will net at least 11 million additional responses, reducing the non-response workload to approximately 40 million. The next step we propose is to use administrative records from trusted federal and state sources to enumerate an additional 6 million housing units. We would then visit the remaining housing units to complete the enumeration. According to our estimates, using administrative records in this way will result in a savings of \$1.4 billion.

We have identified the core set of records we will use for the 2020 Census, but we continue to assess other federal and state sources. However, one of the most important sources, the National Directory of New Hires (NDNH), is currently unavailable for most evaluation and statistical purposes government-wide, including the decennial census, because access to these data for these purposes is not currently permitted under the Social Security Act. This database supplies information on workers, including the newly hired, which the Census Bureau could use to corroborate and supplement other information. Last year, legislation permitting broadened access for Federal evaluation and statistical purposes, including the decennial census, was introduced in and partially moved by Congress. In addition, we developed draft language at the request of the Senate Committee on Homeland Security and Government Affairs. The language provided limited access for the Census Bureau to use these data for statistical purposes and ensured protection of those data under the Census Law. The President's FY 2016 Budget also highlights our need for these data.

Finally, we have reengineered our nonresponse follow-up operation to save an estimated \$2.5 billion. We will now use mobile technology and smartphones to achieve significant efficiencies for the 2020 Census. I should note that we have successfully developed an innovative prototype system that incorporates commercial off-the shelf technology and software to deploy our workforce. We are now able to provide optimized work assignments to our enumerators, including daily route assignments and the best time of day to attempt contact. We are also able to provide the supervisors of our enumerators with real-time updates and alerts regarding the progress of the workers that they oversee.

Supporting these efforts is the Census Enterprise Data Collection and Processing (CEDCaP) effort, which is an enterprise approach to survey and census data collection and processing through shared services. In the past, duplicative systems were created and used for every survey and census. This duplication meant there were more than 100 systems the Census Bureau used to collect and process data, some of which, in the case of the Decennial Census, were decommissioned after use. This stove-piped approach was not the most economical or productive way to do business and the Census Bureau is working toward a sustainable, enterprise approach that simplifies and integrates data systems across the lifecycle from survey design through instrument development, survey, data collection, and data processing, editing, imputation, and estimation.

From dozens of unique systems, we will move to a small suite of shared, reusable systems with the goal of conducting a modernized 2020 Census. Based on my experience, in overseeing the 2000 Census and in the private sector, I am confident we are on course and we have a schedule for the major decisions. Over the course of the last year, we made substantial progress and developed prototype systems to eliminate the paper and pencil processes for the Decennial Census. We established milestones for moving forward and we have engaged Carnegie Mellon's Software Engineering Institute to help us make the critical decisions regarding deployment of production systems.

As a final point, I must mention our commitment to protecting the privacy and confidentiality of individuals' information. The Census Bureau takes cybersecurity seriously. Protecting privacy and confidentiality are central to the Census Bureau's mission and is a core value of our agency—we know that if the public does not trust us to protect their information, they will not provide their information to us. To protect our information systems and the information we collect, the Census Bureau has implemented a robust, comprehensive, and layered cybersecurity program. Cybersecurity is an ongoing process and challenge, and we look to outside experts to constantly evaluate our posture and our security program. Some of the key points of the program are:

- The Census Bureau utilizes the Department of Homeland Security's resources to protect Internet traffic, by looking for malicious code and suspicious activity.
- Other key safeguards include 2-factor authentication; use of encryption in transmissions of data and data at rest; use of a Data Loss Prevention System; use of firewalls, intrusion detection systems and intrusion prevention systems; a dedicated cybersecurity staff that monitors these systems and investigates unusual activity; employee fingerprint and background checks for employees, including enumerators; and all employees receive annual training on data stewardship and security.
- All Census Bureau systems are compliant with the National Institute of Standards and Technology (NIST) Federal Standards and Guidelines, and the Census Bureau security program incorporates continuous monitoring of all IT systems. We have implemented a Risk Management Framework (RMF) program, that is fully compliant with NIST standards and guidelines, and for which the initial step is to identify risks.
- Finally, we work closely with the Department of Commerce Chief Information Officer and Office of Security. We also work with our oversight bodies, the Government Accountability Office (GAO) and Office of the Inspector General, whose recommendations we take seriously.

From CEDCaP and cyber security to the 2016 Census Test and the 2018 Census End-to-End Test, each activity plays a significant role in helping the Census Bureau to design an accurate and cost-effective 2020 Census. The 2018 Census End-to-End Test represents the culmination of research and testing, as we implement the planned census operations in real-time. Even though 2018 may be two years away, we have a sense of urgency right now about what we can accomplish this year to prepare for this test. We cannot risk a major IT system failure in 2020, therefore the 2018 Census End-to-End test is critical and must have all of the major systems in place for this test. The President's Budget includes adequate funding to accomplish our objectives and to conduct a full scale 2016 Census Test, which will provide insights and guide our planning to ensure an accurate census.

The last five years have provided the groundwork for the 21st century census and now, looking forward, we must turn our attention to counting every person in America. But I must also be candid about the next five years and the risks we face. A significant risk we face to executing a 2020 that reduces cost while maintaining quality is receiving adequate funding for the entire lifecycle. Stated simply, if the funding appropriated for each fiscal year during the 2020 Census lifecycle is less than requested, then the Census Bureau must reprioritize its activities, which will affect our ability to reengineer systems and operations to conduct an accurate census. We are fully committed to a census that fairly represents every person in America. We are confident that with the President's Budget request we can build a sustainable census and count every person; however, if we have to defer activities to later years, the costs of an accurate census will escalate and the projected savings will diminish.

I am confident the Census Bureau can achieve these objectives given Congressional support and I look forward to discussing the 2020 Census Operational Plan and other aspects of our planning with you.

Thank you.

Mr. MEADOWS. Thank you so much.
Mr. Cooper.

STATEMENT OF STEVEN I. COOPER

Mr. COOPER. Chairman Meadows, Chairman Hurd, Ranking Members Connolly and Kelly, thank you for the opportunity to testify this afternoon. I am Steve Cooper, Chief Information Officer for the Department of Commerce. It is my pleasure to appear before the subcommittees and update you on our work as we prepare for a successful 2020 Census.

The Census Bureau is working to ensure the necessary information technology, or IT, is in place to support the 2020 Census. The Census Bureau faces an increasing set of challenges, including declining survey participation rates, increased survey costs per household, funding constraints, and cybersecurity threats. These challenges all create risk for the 2020 Census and the Census Bureau has developed a multifaceted strategy that will address these challenges, will transform the business model, and modernize 2020 Census operations.

The 2020 Census requires an IT architecture and infrastructure that is agile, flexible, scalable, and able to accommodate innovations and advances being introduced through IT. As the focus of the 2020 Census shifts from planning to execution, our collective attention must include IT readiness. The Census Bureau has made significant progress in process improvements, IT governance, and closing the skill gaps to ensure it is ready for the 2020 Census.

One of the major enterprise initiatives supporting the 2020 Census is the Census Enterprise Data Collection and Processing initiative, what we call CEDCaP. The CEDCaP initiative aims to create an integrated and standardized suite of systems that will provide shared data collection and processing solutions across all Census Bureau operations. As the Census Bureau proceeds, it is working to ensure that it strikes the right balance between commercial off-the-shelf solutions and custom-developed solutions.

In terms of mobile solutions, the Census Bureau is exploring a mix of Device as a Service and Bring Your Own Device for the 2016 Census Test. The test results will inform subsequent decisions on which mobile devices are leveraged in support of the 2020 Census. In either mobile scenario defined, Census data will be protected at rest and in motion.

The Census Bureau is also exploring cloud computing technology. Various tests are underway to explore processing and storing data in some combination of various cloud infrastructures. For the 2016 Census Test, the Census Bureau will deploy the Internet response option in a secure, FedRAMP-certified, commercially provided private cloud.

The Census Bureau is further exploring how to best employ features like auto-scaling to meet performance demands of the 2020 Census, particularly for systems such as the Internet self-response option that must scale to meet the short-term anticipated demand of millions of users.

In support of this work, the Census Bureau has awarded an initial cloud computing services contract. This contract will also allow

the Census Bureau to gain essential skills and knowledge to be applied during all phases of delivering the 2020 Census.

Obviously, securing confidential data is a major concern for the Census Bureau. The Census Bureau uses an enterprise layered defense strategy to protect its data and systems. For example, the Census Bureau relies on the Department of Homeland Security-managed Einstein program to protect external Internet traffic, and the Census Bureau's internal network is segmented to isolate the systems that are Internet accessible.

The Census Bureau has also worked closely with NIST to implement a risk management framework for all of its systems. Each system also undergoes continuous monitoring to maintain its authorization. This monitoring consists of both automated and manual assessments.

Finally, the Census Bureau continues to work with my office in their implementation of the Department of Homeland Security's Continuous Diagnostics and Mitigation program and our continuing ongoing cyber sprint effort.

Based on my observations to date, the Census Bureau is well positioned to take advantage of early planning, testing, and operational designs. The 2020 Census program is also poised to leverage enterprise initiatives to realize significant efficiencies. However, to adequately implement these strategies and meet the challenges will require the best efforts of the Census Bureau and continued congressional support.

I am deeply grateful for this opportunity to testify before this committee and share these observations, and I'm pleased to answer any questions you may have. Thank you.

[Prepared statement of Mr. Cooper follows:]



**Written Testimony
of
Steven Cooper
Chief Information Officer, Department of Commerce**

**Before the U.S. House of Representatives
Subcommittee on Government Operations
and the
Subcommittee on Information Technology
of the Committee on Government Oversight and Reform**

November 3, 2015

Mr. Chairman and Members of the Subcommittees, I am Steve Cooper, Chief Information Officer for the Department of Commerce.

It is my pleasure to appear before the Subcommittees, and I wish to thank the Chairman and Members for providing me the opportunity to update you on our work for the 2020 Census, and the readiness and security of the information technology at the Census Bureau to effectively support the successful operation of the 2020 Census.

The Census Bureau continues to work to ensure the necessary information technology is in place to support a successful 2020 Census operation and that the critical testing has been conducted or is planned each year leading up to the 2020 Census. This testing ensures the necessary information security controls are in place to protect citizen and household data during data collection and processing.

The Census Bureau faces an increasing set of challenges as it begins preparation for the 2020 Census. Declining survey participation rates, funding constraints, cybersecurity threats, and increased survey costs per household all create risk to the Census Bureau's ability to deliver high quality, timely, relevant, and cost-effective information. To address these risks, the Census Bureau has developed a multi-faceted change strategy that will address these challenges, transform the business model, and modernize 2020 Census operations. The modernized 2020 Census requires an information technology architecture and infrastructure that is agile, flexible, scalable, and able to accommodate innovations and advances in technology.

Effective application of technology plays a key role in the Bureau's strategy and enables the rethinking and reengineering of existing business processes. Technology also serves as a change catalyst for developing enterprise capabilities that will create new architectures and modernize data and systems management for the 2020 Census and all Census Bureau censuses and surveys— from data collection and processing to data dissemination.

Since 2009, the Department of Commerce and the Census Bureau have focused on the application of lessons learned and fostering of strategic alignment with Government IT initiatives beginning with the 25 Point Implementation Plan to Reform Federal Information Technology (IT) Management. The Census Bureau continues to align its IT goals with subsequent initiatives, including those found in Memorandum M-11-29, Chief Information Officer Authorities: Memorandum M-12-10, Implementing PortfolioStat; and now the Federal IT Acquisition Reform Act (FITARA) in support of Memorandum M-15-14, Management and Oversight of Federal IT. Common themes in all of these directives, and thereby the Census Bureau's IT strategies, include IT transformation, improving operating efficiencies, leveraging shared services, and increasing Chief Information Officer (CIO) accountability.

More recently, the Census Bureau has made a number of its major design decisions for the 2020 Census. These decisions have been documented in the 2020 Census Operational Plan, which the Census Bureau issued last month – three years earlier than for the 2010 Census, and is grounded in the Census Bureau's Enterprise IT Strategy. The effort included developing the operating plan for the IT Infrastructure Operation and Decennial Service Center Operation, as well as reviewing all of the IT capabilities required by the remaining operations identified in the 2020 Census Operational Plan.

As the focus of the 2020 Census Program shifts from planning to execution, our collective attention must include IT Readiness for the 2020 Census. The Information Technology Directorate, under leadership of the Acting Census Bureau CIO, continues to align the enterprise IT Roadmap with the 2020 Operational Plan to ensure IT readiness for the 2020 Census, as well as all of the testing leading up to the 2020 Census. The IT Directorate has defined and provided enterprise guidelines to evaluate and ensure systems readiness, and has imbedded a Chief Solutions Architect, Chief Program Engineer, Chief Security Architect and other systems engineers within the 2020 Systems Engineering and Integration (SE&I) Team.

The Census Bureau also has made significant progress in process improvements, IT governance, and closing the systems engineering IT skills gap to ensure IT Readiness for the 2020 Census. This progress includes:

- Implementing and improving standard processes and support resources for portfolio management, program management, and project management;
- Establishing and updating the Enterprise Systems Development Lifecycle (ESDLC) to include agile software development in addition to the traditional waterfall method;

- Establishing the Office of Systems Engineering that has defined standard engineering processes, artifacts and reviews aligned to the program-level and project-level phases and governance process;
- Increasing expertise and centralized support resources in the areas of project management, schedule management, requirements development and management, solution design and analysis of alternatives, and testing; and
- Working closely with the Strategic Workforce Planning Initiative to ensure that proper IT skills are available when needed as identified by the various Census Bureau directorates.

To support these efforts, the Census Bureau engaged General Service Administration's 18F experts about the use of Schedule A hires for 2-year term appointments that fill critical gaps in technical areas, including cloud computing, and security.

In order to streamline the acquisition of needed contract resources, the Census Bureau successfully completed contract awards for the Systems Engineering & Integration Enterprise Solutions Framework (ESF) for use by all directorates acquiring SE&I services and solutions. This allows for a standard and streamlined process for acquiring SE&I services with up to a \$2.5B IT Investment Authority through 2022. Other Department of Commerce (DOC) Bureaus can leverage the SE&I ESF and several bureaus have already expressed interest in doing so. The Census Bureau has already awarded Initial Work Orders.

One of the major enterprise initiatives providing support to the 2020 Census is the Census Enterprise Data Collection and Processing (CEDCaP) initiative. The CEDCaP initiative aims to create an integrated and standardized system of systems that will offer shared data collection and processing across all Census Bureau operations. CEDCaP will yield the following benefits:

- Creation of an enterprise-wide data collection shared service that provides the functional capabilities to meet requirements across the Census Bureau, including the Decennial Census, under a single, integrated and centrally managed program;
- Ability to reengineer current business processes, such as the implementation of an adaptive design methodology, across all surveys and censuses to increase efficiencies in data collection and processing;
- Realization of cost savings through the prevention of the initiation of single-use survey specific systems and retirement of unique, survey-specific systems and redundant capabilities;
- Reduction of risks for the 2020 Census by providing an integrated, proven solution well in advance of the Decennial Census; and the
- Realization of the Census Bureau's Information Technology Guiding Principles to simplify, innovate, and engage by looking to the cloud first and emphasizing standard-based, commercial off-the-shelf (COTS) solutions over custom development.

In terms of governance, the CEDCaP program achieved the DOC Milestone 2 Review Approval from the Deputy Secretary of Commerce in July 2015. The CEDCaP Program follows the Program LifeCycle (PgLC) Framework, based on DOC's Acquisition Framework, and Enterprise Systems Development LifeCycle to manage program requirements, development, and execution. Program phase gate reviews are approved by the Program LifeCycle Investment Review Board (PLIRB) governance process to include phase gate reviews. The Census IT Directorate Project Review (CIPR) board approves project phase gate reviews.

To ensure that the agency strikes the right balance between COTS and custom-developed solutions, the Census Bureau is comparing in-house custom solutions with COTS solutions specifically focusing on internet and mobile data collection, dynamic case management, and dash boarding to finalize build/buy decisions for the enterprise and 2020 Census, scheduled for October 2016.

Carnegie Mellon's Software Engineering Institute is independently evaluating the CEDCaP program's process for analyzing and assessing build versus buy decisions and performing an independent assessment of the custom and COTS products, also scheduled for October 2016.

In terms of mobile solutions, the Census Bureau is exploring the technology it will employ to enable data collection in the field and will be using a mix of Device as a Service (DaaS) and Bring Your Own Device (BYOD) for the major 2016 in-field test. The Census Bureau will award a DaaS contract to provide hardware, wireless service, and accessories. The vendor will be responsible for ensuring adequacy of the wireless service coverage employed in a given location. Although the Census Bureau has not made a final determination on specific make/model/form factor of data collection devices for the 2020 Census, the Census Bureau is developing field data collection applications to work on multiple form factors (tablet/smartphone) and multiple operating systems (Android/iOS) to be flexible and allow for BYOD use. The results of the 2016 Census Test will inform subsequent decisions on which mobile devices are leveraged in support of the 2020 Census.

In both scenarios, census data will be protected "at rest" and "in motion." The Census Bureau will use FIPS 140-2 encryption algorithms for the data "at rest" and a combination of the Mobile Device Management (MDM), Mobile Application Management (MAM), and virtual private network (VPN) protections for "in motion." In addition, the Census Bureau further protects the application with a username and password, required each time a Census Bureau field staff accesses the applications. The Census Bureau is working on secure implementation of derived credentials, which would provide for multi-factor identification based on PIV.

Another key technology the Census Bureau is exploring is Cloud Computing. The Census Bureau has approval through its Data Stewardship Executive Policy Committee to host confidential census data in FedRAMP certified clouds that meet legal and security requirements and is exploring the process and of storing other titled data (e.g. Title 26) in the cloud. Various tests are underway to process and store data both in FedRAMP certified commercially-provided private clouds, a virtualized Census Bureau private cloud, and potential use of hybrid cloud infrastructures. In the 2016 Census Test, the intent is to deploy the internet data collection instrument in a FedRAMP-certified commercially provided private cloud.

The Census Bureau adheres to all NIST requirements and guidance as required under the Federal Information Security Management Act (FISMA). For security for the Census Bureau's cloud initiative the Census Bureau relies on the GSA FedRAMP process, as well as mapping the security information provided by the cloud vendors to the Census Bureau's internal Risk Management Framework security profile so that the CIO can have a consistent understanding of the security posture of all IT systems supporting the Census Bureau mission. Testing and experimentation in cloud deployments began in Summer 2015 using Amazon Web Services to include taking advantage of auto-scaling to the provision of computing resources to match the demand required to meet specified performance at minimum cost. The 2016 Census Test will take advantage of FedRAMP Certified Commercial Cloud services and will provide cost basis estimates as well as allow the Census Bureau to acquire hands-on learning in migration, integration, operation, monitoring, and maintenance of services within a cloud model.

Census will determine how best to employ features like auto-scaling to meet performance demands of the 2020 Census, particularly for systems such as the internet self-response capability that must scale to meet the short-term anticipated demand of millions of users, while containing costs. The Census will acquire on-demand infrastructure, to the greatest extent possible, through the FedRAMP cloud. The goal of the Census Bureau is to satisfy the scale and resources required by the Census Bureau while minimizing the impact on its own internal footprint. In continuing support of this work, the Census Bureau has awarded a Cloud Enterprise Solution Framework for Cloud Computing Services contract, established as the initial move into cloud computing. This contract allows the Census Bureau to begin migration of services into FedRAMP-certified commercial clouds and gain essential skills and knowledge.

Obviously, security of confidential data is a major concern for the 2020 Census, as well for all of the Census Bureau's other censuses and surveys. The Census Bureau uses an enterprise layered defense strategy to protect its data and systems. They are protected by the DHS managed Einstein program. In addition to this, the Census Bureau uses a cloud service provider to provide protection from Distributed Denial of Service (DDoS) attacks. The Census Bureau also has security at the perimeter of the network through routers and firewalls and further has its own Intrusion Detection Systems and Intrusion Prevention Systems (IDS/IPS) in place. They have a segmented network to isolate the internal network from systems that are Internet accessible.

Further, they have implemented a Risk Management Framework (RMF), based on NIST Special Publication 800-37 rev.1. The Census Bureau worked with NIST in the development of the RMF, and the resulting framework has been recognized by NIST as being very solid and meeting all the intentions of Publication 800-37 rev1. All of the Census' systems are fully assessed and have current Authorizations to Operate. Each of them undergoes continuous monitoring to maintain the authorization. This monitoring consists of both automated and manual assessments. Finally, the Census Bureau, in alignment with the DOC, has implemented the Department of Homeland Security (DHS) Continuous Diagnostics and Mitigation (CDM) tools under Phase 1 and is currently working with the Department and DHS to implement additional tools under Phase 2c of the CDM program. The Census Bureau has also reviewed all available information concerning the OPM hack and is making changes to our program, as needed, based on what we have learned. The Census Bureau has also worked closely with DOC on the on-going cyber sprint effort and strengthening our entire IT Security Program, which includes support for the 2020 Census.

The Census Bureau continues to work with the Government Accountability Office (GAO) to close all the findings of the 2013 audit report. As of October 19, 2015, GAO has closed 65 of the 115 recommendations and referred 19 back to the Census Bureau for additional information. Thirty-one of our original responses are still pending review at GAO. The Census Bureau communicates with GAO on a bi-weekly basis to provide the additional information as the Census gathers it and to check on the status of additional GAO findings.

Finally, during this period of IT leadership transition, the Census Bureau has continued to effectively and securely deliver the solutions needed across all directorates, achieving and even exceeding its IT goals, including:

- Effective support for the Decennial 2015 National Content Test and preparing for the Field IT Infrastructure in support of the 2016 Census Test;
- Timely delivery of Census Enterprise Data Collection and Processing (CEDCaP) Product Sub-Releases and Releases and associated system components within budget to support the production needs of the Demographic, Economic, and Decennial directorates;
- Successful contract awards for initial Cloud Services Enterprise Solutions Framework (ESF) to five (5) Service Disabled, Veteran-Owned Businesses (SDVOB);
- Alignment with the Federal Information Technology Acquisition Reform Act (FITARA); and
- Establishment of, and update to, the Enterprise Systems Development Lifecycle (ESDLC) to support agile software development methodologies.

Based on my observations, the Census Bureau is well positioned as compared with prior decades to take advantage of early planning, testing, and operational designs. The 2020 Census program is also well positioned to leverage enterprise initiatives to realize significant efficiencies. Innovations in cloud computing and mobile technologies show great promise for enhancing IT-based solutions in support of the 2020 Census. To adequately implement these strategies and meet the challenges listed above will require the best efforts of the Census Bureau and continued Congressional support.

I am deeply grateful for this opportunity to testify before this committee and share these observations, and I am pleased to answer any questions you may have.

Mr. MEADOWS. Thank you, Mr. Cooper.
Mr. Goldenkoff.

STATEMENT OF ROBERT GOLDENKOFF

Mr. GOLDENKOFF. Chairmen Meadows and Hurd, Ranking Members Connolly and Kelly, and members of the Government Operations and IT Subcommittees, thank you for the opportunity to be here today to discuss the progress the Census Bureau is making in controlling the cost of the 2020 enumeration.

The Bureau's goal is to conduct the Census at a lower cost per household than the 2010 Census, adjusted for inflation, while maintaining accuracy. This would mean less than around \$94 per household. It's an extremely difficult task as the population is growing steadily larger, more diverse, increasingly difficult to find, and less willing to participate in the head count. Moreover, the Census is conducted on a tight schedule with little room, if any, for slippage.

The Bureau plans to hold down costs in part by expanding its use of data that has already been collected by other government agencies in the course of administering their programs. Known as administrative records, such information can help improve accuracy and reduce the need for labor-intensive field operations, especially during one of the most expensive of all Census activities, Nonresponse Followup. The Bureau has used administrative records in previous decennials, but not to the same extent as that planned for 2020.

In my remarks today, which are based on a report we issued last month, I will describe the opportunities and challenges the Bureau faces in using administrative records and the steps it needs to take going forward to help ensure they produce the desired results.

The Bureau estimates it can save up to \$1.4 billion compared to traditional census-taking methods by using administrative records for three purposes during Nonresponse Followup. They include identifying and removing vacant and nonexistent housing units from the follow-up workload before Census workers start knocking on doors. Second, enumerating non responding occupied housing units if the information meets a certain quality threshold. And third, predicting the best times to visit a household.

Using administrative records for these activities can greatly improve productivity. For example, in a test conducted in Arizona's Maricopa County earlier this year, the Bureau reduced the follow-up workload by 11 percent by removing vacant and nonexistent households.

The Bureau is also exploring nine additional applications of administrative records that may help reduce cost or improve quality still further. The Bureau currently has access to data held by the U.S. Postal Service, IRS, and Selective Service System, among other agencies. It's also considering other data sets, such as the National Directory of New Hires. The Bureau believes the NDNH and other records could improve its ability to find historically hard-to-count populations, such as certain minority groups and young children. However, the Bureau still needs to secure statutory access to the NDNH.

The Bureau's planned use of administrative records is commendable, but much work remains. For example, in our October report we recommended that the Bureau set deadlines to decide which, if any, of the nine additional uses of administrative records still under consideration will be used in 2020. This will help ensure the Bureau has sufficient time to review the data sets, determine their fitness for use, and fully test them.

Final decisions are needed by the end of fiscal year 2017 in order to be included in the Bureau's full end-to-end test in 2018. However, these deadlines do not appear in schedule documents.

The Bureau must also continue to address challenges to using administrative records, such as protecting confidential information and ensuring congressional and public acceptance of the Bureau's plan to share personal data across government agencies. Fully implementing our prior recommendations to strengthen the security of its information systems and developing a congressional outreach strategy could help address these challenges.

In summary, the bureau has made noteworthy progress in expanding its use of administrative records. Going forward, though, perhaps one of the biggest risks the Bureau faces is the ticking clock. Any delays could have serious implications for downstream activities. As a result, continued congressional attention will be needed to help ensure that the Bureau, one, stays on schedule; two, sets deadlines for key go/no-go decisions on additional uses of administrative data; and three, fully implements our prior recommendations.

This concludes my remarks, I'll be happy to answer any questions that you may have.

[Prepared statement of Mr. Goldenkoff follows:]

GAO Highlights

Highlights of GAO-16-206T, a testimony before the Subcommittees on Government Operations and Information Technology, Committee on Oversight and Government Reform, House of Representatives

Why GAO Did This Study

With a life-cycle cost of about \$13 billion, the 2010 Census was the most expensive U.S. census in history and was 56 percent more costly than the \$8.1 billion 2000 Census (in constant 2010 dollars). The Bureau estimates that its use of administrative records in the 2020 Census will reduce the cost compared to traditional census methods by \$1.4 billion.

Given the potential cost savings associated with the use of administrative records, this testimony, which is based on a report GAO released last month, focuses on (1) the Bureau's plans for using administrative records, and the opportunities and challenges the Bureau faces going forward; and (2) the key assumptions supporting estimates of expected cost savings. To meet these objectives, GAO reviewed Bureau planning documents and test plans, and interviewed Bureau officials. GAO also relied on its Schedule Assessment Guide.

What GAO Recommends

In its report issued last month GAO recommended that the Census Director set deadlines for making final decisions about which records to use, and for what purpose. This will help ensure the Bureau's resources focus on those activities that show the most promise for reducing enumeration costs. The Department of Commerce—the Bureau's parent agency—concurred with GAO's findings and recommendation.

View GAO-16-206T. For more information, contact Robert Goldenkoff at (202) 512-2757 or goldenkoffr@gao.gov.

November 3, 2015

2020 CENSUS

Progress Report on Using Administrative Records to Control Enumeration Costs

What GAO Found

The U.S. Census Bureau (Bureau) estimates that it can save around \$1.4 billion using administrative records, compared to relying solely on traditional enumeration methods. While the Bureau has made some limited use of administrative records during past decennials, it plans to use them much more extensively in 2020 to achieve these savings.

For example, the Bureau plans to use administrative records to reduce the field work required for its most expensive census operation—nonresponse follow-up—when temporary Bureau employees knock on doors across the country to obtain information from people who did not respond to the census, or who were missed by census mailings. According to the Bureau, using administrative records to (1) identify vacant housing units; (2) identify and enumerate occupied nonresponding housing units when the records meet a certain quality threshold; and (3) predict the best times to visit a household can generate substantial cost savings. The Bureau is also exploring the feasibility of nine additional uses of administrative records that could reduce costs and improve the quality of the census still further. The Bureau already has access to nearly all of the data sources it needs to achieve the desired cost savings. It is also working to gain access to additional databases that could help improve its ability to find historically hard-to-count populations, such as certain minority groups and young children. While the Bureau is to be commended for its efforts to expand its use of administrative records, going forward, it will be important to set deadlines to help ensure it makes timely decisions on these other databases and uses of administrative records. According to Bureau officials, final decisions on the use of administrative records are needed by the end of fiscal year 2017 so the records can be adequately tested in the Bureau's full end-to-end test in 2018. However, these deadlines do not appear in schedule documents.

It will also be important for the Bureau to address key challenges to using administrative records, including (1) ensuring the quality of the records it receives from other government agencies; (2) protecting confidential data; and (3) ensuring congressional and public acceptance of the Bureau's plan to share personal data across government agencies. The Bureau's ongoing research and testing efforts can help with the first challenge. Fully implementing our prior recommendations to strengthen the security of its information systems and to develop a congressional outreach strategy could help address the second and third challenges.

Key assumptions the Bureau used in estimating potential cost savings from administrative records are logical, and the Bureau plans to provide additional support for them. For example, the Bureau's assumption that it could reduce its follow-up workload follows clearly from the Bureau's use of administrative records to remove vacant units from among those housing units needing follow-up because people did not respond to the census, reducing that workload by 11.6 percent. The Bureau released an updated life-cycle cost estimate in October 2015, and GAO anticipates reviewing its reliability after the Bureau makes support for the estimate available.

Chairmen Meadows and Hurd, Ranking Members Connolly and Kelly, and Members of the Subcommittees:

I am pleased to be here today to discuss the U.S. Census Bureau's (Bureau) efforts to control the cost of the 2020 Census by making greater use of information already provided to federal and state governments as they administer other programs—what is referred to as “administrative records.” With a life-cycle cost of about \$13 billion, the 2010 Census was the most expensive U.S. census in history and was 56 percent more costly than the \$8.1 billion 2000 Census (in constant 2010 dollars). Given budgetary realities, that cost growth is simply unsustainable. Beginning in 1990, we reported that rising costs and difficulties in securing public participation, among other challenges, required a new approach to taking the census.¹ Part of the Bureau's response is its planned use of administrative records. The Bureau estimates their use could generate as much as \$1.4 billion in cost savings compared to traditional census taking methods.

Today's hearing is timely, as our past reviews of prior decennials have underscored the importance of (1) early and ongoing congressional oversight to keep census preparations on track, and (2) congressional acceptance of the Bureau's overall approach. From the very beginning of the 2020 life cycle, we have been reporting on the Bureau's progress in re-engineering the census, including its planned design and operational changes, as well as its Information Technology (IT) management, the latter of which my colleague is addressing in her statement.² In my remarks today, I will describe the opportunities and challenges the Bureau faces in using administrative records as well as the key assumptions supporting the Bureau's related estimated cost savings.

My testimony is based on our review of the Bureau's plans to use administrative records in the 2020 Census that we issued last month, and includes an update to some estimated cost data obtained from the

¹See for example, GAO, *2000 Census: Progress Made on Design, but Risks Remain*, GAO/IGD-97-142 (Washington, D.C.: July 14, 1997), and *Decennial Census: Preliminary 1990 Lessons Learned Indicate Need to Rethink Census Approach*, GAO/T-IGD-90-18 (Washington, D.C.: Aug. 8, 1990).

²GAO, *2020 Census: Key Information Technology Decisions Must be Made Soon*, GAO-16-205T (Washington, D.C.: Nov. 3, 2015).

Bureau.³ For that work, we reviewed Bureau planning documents and test plans, our prior reporting on decennial census planning, interviewed Bureau officials, and relied on our Schedule Assessment Guide.⁴ We did not assess the reliability of the Bureau's estimated cost-savings discussed in today's statement because the Bureau released an update in early October, which we will be reviewing soon. The work on which this statement is based was conducted in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

We, the Bureau, and others have observed that some of the information the Bureau collects during the census has already been gathered by other government agencies in the course of administering their programs. Accessing that information could provide the Bureau with data to help conduct the census and, in some cases, complete census forms that have missing data. Such uses of administrative records have the potential to reduce the cost of the decennial census because, for example, the Bureau would need to hire fewer temporary workers and acquire less office space and equipment to support fieldwork. Moreover, some of the information collected through administrative records could be more accurate than information the Bureau collects through traditional door-to-door follow-up methods, such as when the Bureau's enumerators need to interview neighbors or other "proxy" respondents because they cannot reach a household member to collect needed information.

While the Bureau has made some limited use of administrative records during past decennials, it plans to use them much more extensively in 2020. In its first operational plan for the 2020 Census (released October 6, 2015), the Bureau reported design decisions including the use of administrative records to identify vacant addresses in advance of follow-up field work and to enumerate nonresponding households when possible

³GAO, *2020 Census: Additional Actions Would Help the Bureau Realize Potential Administrative Records Cost Savings*, GAO-16-48 (Washington, D.C.: Oct. 20, 2015).

⁴GAO, *GAO Schedule Assessment Guide: Best Practices for Project Schedules—Exposure Draft*, GAO-12-120G (Washington, D.C.: May 30, 2012).

in order to reduce the need for repeated contact attempts during its nonresponse follow-up operation (NRFU). The Bureau also updated the life-cycle cost estimate for the 2020 Census to \$12.5 billion (in constant 2020 dollars).⁵ This is slightly lower than the Bureau's prior life-cycle cost estimate of \$12.7 billion. We plan to assess the reliability of the new cost estimate and examine the practices the Bureau used to produce it after the Bureau makes the model and its supporting documentation available. As part of the operational plan, the Bureau also released an update of how much its fundamental redesign of the census would reduce costs in four major design areas compared to the cost of a 2020 Census conducted using the methods of the 2010 Census. Table 1 below compares the cost reduction in those four areas.

Table 1: The Census Bureau's Cost Estimates Show Redesign to Cost Less Than Traditional Census

Key design area	August 2014 estimated savings	October 2015 estimated savings
Reengineering address canvassing	\$1 billion	\$900 million
Optimizing self-response	\$548 million	400 million
Using administrative records	\$1.2 billion	1.4 billion
Reengineering field operations	\$2.3 billion	2.5 billion
Total savings compared to Bureau's projected cost of 2020 Census using traditional approach and methods (in 2020 constant dollars)	\$5.1 billion	\$5.2 billion

Source: GAO analysis of Census Bureau data. | GAO-16-206T

The operational plan describes over 100 other preliminary design decisions related to the Bureau's efforts to build an address list and collect census responses. Its release marks a critical turning point in the decade-long countdown to the next decennial, as the Bureau completes its early research and testing plan, and pivots toward developing operations and systems and testing them to refine the census design.

⁵In GAO-16-48 we reported a different number because the paper version of the Bureau's Operational Plan for the 2020 Census originally provided to us, reported a life-cycle cost estimate of \$12.3 billion, which was in current dollars.

**Administrative
Records Can Reduce
2020 Fieldwork, but
the Bureau Will Need
to Better Define
Milestones and
Deadlines to Help
Manage Risks**

Bureau officials have said they hope to use administrative records to reduce the field work involved in the most expensive census operation—NRFU, when Bureau staff traditionally knock on doors across the country at homes of people who did not respond to the census, or who were missed by census mailings. The Bureau has reported that the following three uses are key to potentially saving up to \$1.4 billion compared to using traditional census methods.

Identify vacant housing units. The Bureau incurs a large part of its census cost while following up at residences that did not return a census questionnaire. However, during the 2010 Census, of the 48 million housing units enumerators visited for follow up, about 14 million were vacant. One of the largest efficiency gains to the census may come simply from using administrative records to remove these vacant units from the follow-up workload. In a test in Arizona earlier this year, this use of administrative records enabled the Bureau to reduce the NRFU workload by 11 percent.⁶ Since we completed our audit work, the Bureau announced that it would still send a reminder post card to units it identified as vacant, which will cost more, but will provide any missed household one more opportunity to respond to the census.

Identify and enumerate occupied nonresponding housing units. During the 2015 Census Test, the Bureau demonstrated it could use administrative records to accurately count some nonresponding occupied households if the household had administrative records meeting a certain quality threshold, without attempting any visits. In the test, the Bureau used this approach to reduce the NRFU workload by about 20 percent. Since we completed our audit work, the Bureau announced that before using administrative records to enumerate such households in 2020, it will still attempt one visit to the household. Attempting a visit will add costs, but will also provide the household one more opportunity to respond to the census.

Predict the best times to complete NRFU. One of the challenges the Bureau faces when knocking on doors is reaching a household when someone is home. In the 2015 Census Test, the Bureau used

⁶Earlier this year in Maricopa County, Arizona, the Bureau conducted its 2015 Census Test to see how well it can use administrative records to reduce fieldwork and increase productivity for NRFU. The test also included a new field management structure and an enhanced Operations Control System supporting daily reassignments of cases.

demographic information, such as age, from administrative records sources in addition to information about how households had responded to other Bureau surveys in order to help determine the contact strategy for deciding if and when to interview a household.

The Bureau tested each of these uses during its 2015 Census Test and plans further testing of them to refine the methods, but the Bureau has already decided to use them.

The Bureau Plans Additional Opportunities for Using Administrative Records in the 2020 Census

In addition to the three uses the Bureau has committed to, the Bureau has identified nine additional uses of administrative records that may help further reduce cost or improve the quality of the census (see figure 1). The Bureau has not separately estimated cost savings for these nine uses, but has begun researching the feasibility of most of them.

Figure 1: The Census Bureau Has Identified Nine Additional Opportunities for Administrative Records in 2020 Census Activities

2020 Census Activity (as of August 2015)	When activity occurs relative to data collection		
	Before	During	After
Validate and update address list	✓		
Identify and target outreach to group quarters ^a	✓	✓	
Quality control		✓	
Non-ID processing – Locate respondents' addresses ^b		✓	
Non-ID processing – Validate responses ^b		✓	
Impute occupancy status for unresolved households			✓
• Impute count of previously unresolved households			✓
• Impute age/sex/race/ethnicity of previously unresolved households			✓
Evaluate Census accuracy			✓

Source: U.S. Census Bureau. | GAO-16-206T

^aGroup quarters consist of college dormitories, prisons, nursing homes, and other facilities typically owned or managed by an entity providing housing, services, or both for the residents.

^bNon-ID processing accommodates certain self-responses the Bureau may receive from households that may have lost or never received mailings or other advance communication with an ID number from the Bureau, determining where to count them (locate them) and attempting to reduce their duplication and falsification (validate the responses).

As shown in the figure, these uses would occur during various points relative to data collection.

Before data collection. The Bureau is already using administrative records to validate and update the address list. The Bureau is drawing on address lists and map information from state, local, and tribal governments, in addition to information obtained from commercial sources, to update its own address list continuously throughout the decade, reducing the need for a more costly door-to-door canvassing during the 2 years prior to the census, as was done for the 2010 Census. In addition, Bureau officials reported research is about to begin on how to better use records to identify group quarters, such as dormitories, prisons, nursing homes, and homeless shelters, and to target outreach, that is, encourage cooperation of staff at these locations with the census. The Bureau historically uses special procedures to enumerate at these places, and administrative records could potentially jump start the time and effort spent getting ready for them.

During data collection. The Bureau is considering using administrative records in lieu of some follow-up visits for the purpose of quality control of field work. In past decennials, the Bureau has called or sent enumerators to re-interview some respondents. Relying on administrative records could reduce fieldwork and respondent burden, or enable the Bureau to better target re-interviews of respondents. The Bureau is also researching how administrative records can be used to help process responses that do not have a census ID number on them (this activity is called non-ID processing). The Bureau may receive such responses from households that lost or never received mailings or other advance communication with an ID number from the Bureau. A test in 2015 in the Savannah, Georgia, media market area demonstrated that a large collation of administrative records from many sources was effective in helping the Bureau correct or fill in missing address information. This enabled the Bureau to better locate where those responses should be counted. The Bureau is also researching how other records may help it validate responses or the identities of those who submit responses as part of this processing.

After data collection. When the Bureau still does not have information on a housing unit after collecting data during field operations, it will attempt to impute the data—as it has done since 1970. According to Bureau officials, in 2020 the Bureau will use administrative records to help improve how it imputes three related types of data: (1) whether or not a unit is occupied, (2) how many people live in the unit, and (3) the residents' demographic characteristics, such as sex, race, and ethnicity. Finally, the Bureau is considering how administrative records might help it evaluate census accuracy.

The Bureau Has Identified Potential Sources for Most of the Administrative Records It Hopes to Use for the 2020 Census

The Bureau has identified and obtained access to nearly all of the sources that it believes it needs in order to leverage all of the opportunities it has identified (see figure 2).

Figure 2: The Census Bureau Obtained Access to and Tested Sources of Administrative Records

Sources	Access obtained and fitness screened	Testing ongoing
United States Postal Service – Delivery Sequence File, Undeliverable As Addressed (UAAs)	✓	✓
Centers for Medicare & Medicaid Services – Medicare enrollment	✓	✓
Indian Health Service – Patient registration	✓	✓
Internal Revenue Service – Individual 1040s and informational returns 1099s	✓	✓
Social Security Administration – Number identification file	✓	✓
Housing and Urban Development – Public and Indian Housing Center data; Tenant Rental Assistance Certification System; Computerized Homes Underwriting Management System	✓	✓
Selective Service System – Registration file	✓	✓
Health and Human Services – National Directory of New Hires		
Social Security Administration – KidLink		
States – Temporary Assistance for Needy Families (TANF); Supplemental Nutrition Assistance Program (SNAP); Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)	(9 states)	✓
State, local and tribal governments – Addresses and maps	(1,000 localities)	✓
Commercial sources	✓	✓

Source: GAO analysis of Census Bureau information. | GAO-16-206T

In addition to the records already obtained, the Bureau is working to gain access to the **National Directory of New Hires (NDNH)**, a national database of wage and employment information used for child support enforcement, and **KidLink**, a database from the Social Security Administration that links parent and child Social Security numbers for children born after 1998 in U.S. hospitals. Both of these databases would help the Bureau improve its ability to find historically "hard-to-count" groups, such as certain minority groups or young children.

Bureau officials stated that they are examining ways to quantify the potential effect that their access to these additional sources could have

on the 2020 Census. But they point out that there is value in accessing these records for the Bureau's other statistical surveys as well, and that even if they are unable to obtain the additional records in time for the 2020 Census, they would continue pursuing them for these other purposes, as well as for use in future censuses.

Setting Deadlines Would Help the Bureau Ensure It Makes Timely Decisions on Administrative Records

While the Bureau is to be commended for its efforts to expand its use of administrative records to control costs and increase accuracy, we identified actions the Bureau could take to increase its chances of success.⁷ First, as of August 2015, the Bureau had not set deadlines to determine which of its identified uses of administrative records it will or will not implement for the 2020 Census, nor had it set deadlines for determining exactly which records from which sources it will tap in support of each use it implements. Moreover, the Bureau had no deadlines against which to measure progress for obtaining access to additional sources or scheduled milestones for when key steps may need to be taken in order to integrate them within 2020 preparations. For example, the Bureau will need time to review files in order to ensure their fitness for use before the Bureau can integrate them into the census design. Earlier in October the Bureau announced time frames for several decisions related to uses of administrative records and explained that the decisions are not yet included in its integrated schedule of activities. According to our scheduling guide, assurance of program success can be increased when management relies on credible schedules containing the complete scope of activities necessary to achieve established program objectives.⁸

Bureau officials have stated that final decisions on the use of administrative records are needed by the end of fiscal year 2017 in order to be included in the Bureau's 2018 full end-to-end test. However, these deadlines do not appear in schedule documents. We recommended that the Census Director ensure that resources focus on activities with promise to reduce cost by documenting milestones related to deciding which records to use, particularly for purposes not yet demonstrated as

⁷GAO, *Information Security: Actions Needed by Census Bureau to Address Weaknesses*, GAO-13-63 (Washington, D.C.: Jan. 22, 2013). Another version of this report was issued for limited distribution.

⁸GAO-12-120G.

feasible or involving records it does not already have access to, such as NDNH and KidLink. The Department of Commerce—the Bureau's parent agency—concurred with our recommendation. Deadlines for deciding on all potential uses—either committing to move forward with them or abandoning them as possibilities for 2020—and for deciding how all other records will be used would help to ensure the Bureau is using its resources cost-effectively.

The Bureau Is Taking Steps to Address Challenges to Using Administrative Records for the 2020 Census; Implementing Our Previous Recommendations Could Help

The Bureau is taking steps to address challenges it faces in using administrative records to control costs and improve the quality of the 2020 Census.

One challenge facing the Bureau is ensuring the quality of the records it receives from other agencies and levels of government. To meet this challenge, the Bureau has processes in place and is conducting research and testing to ensure quality of records. For example, to ensure accuracy, the Bureau routinely screens address and map files provided by state, local, and tribal governments to determine if they satisfy preset minimum quality standards for completeness of address information. This helps to improve the master list of addresses. The Bureau plans comprehensive testing of all records during an end-to-end test of its 2020 Census design (to be conducted in 2018). The Bureau plans additional testing of administrative records for the 2016 Census Test in the Los Angeles and Houston metro areas, a large test of address canvassing for 2016, and an additional site test in 2017 at an undetermined location.

The Bureau reported it will review imputation models it used during prior censuses to determine how it can integrate information from administrative records into them in fiscal year 2016. Tests will be included in the 2016 Census Test.

A second challenge involves protecting confidential data. We have previously reported that until the Bureau implements a complete and comprehensive security program, it will have limited assurance that its information and systems are being adequately protected against unauthorized access, use, disclosure, modification, disruption, or loss. In January 2013, we made over 100 recommendations aimed at addressing

weaknesses in that program.⁹ The Bureau agreed and Bureau officials state that the Bureau has taken action on all 115 of our recommendations to improve its security program. In assessing the Bureau's reported actions, we have reviewed documentation pertaining to 97 of the recommendations—66 of which we have confirmed have been addressed and 31 that require additional actions and/or documentation from the Bureau. We are currently analyzing the extent to which the remaining 18 recommendations have been addressed by the Bureau and expect to complete that review by the end of 2015. My colleague is further addressing the security challenge in her statement.¹⁰ Bureau officials pointed out that the Bureau is well positioned to prevent disclosure of administrative records, as it has long-standing experience in collecting data from other agencies and reporting on them. Furthermore, the Bureau and the agency providing the data agree to data safeguards during negotiations for access.

A third challenge concerns public acceptance and attitudes about sharing of personal data across government agencies for the purposes of the census. We have previously reported on the need within the federal statistical system for broader public discussion on balancing trade-offs among competing values, such as quality, cost, timeliness, privacy, and confidentiality.¹¹ The public has related concerns involving trust in the government and perceptions about the burden on respondents as well the social benefits of agencies sharing data. We recommended in 2012 that the Bureau develop and implement an effective congressional outreach strategy, particularly on new design elements the Bureau is researching and considering as well the cost-quality trade-offs of potential design decisions.¹² The Bureau agreed with the recommendation and, in November 2014, it provided us with a congressional engagement plan. The four-page plan brings together in one place a summary of the Bureau's ongoing activity in this area, yet, by itself, lacks goals or

⁹GAO, *Information Security: Actions Needed by Census Bureau to Address Weaknesses*, GAO-13-63 (Washington, D.C.: Jan. 22, 2013).

¹⁰GAO-16-205T.

¹¹GAO, *Federal Statistical System: Agencies Can Make Greater Use of Existing Data, but Continued Progress Is Needed on Access and Quality Issues*, GAO-12-54 (Washington, D.C.: Feb. 24, 2012).

¹²GAO, *2020 Census: Additional Steps Are Needed to Build on Early Planning*, GAO-12-626 (Washington, D.C.: May 17, 2012).

strategies for attaining them, or accountability for who will work to implement them or when. We will continue monitoring the Bureau's efforts to address this recommendation, particularly as these efforts may depend on scheduling of activities the Bureau may yet set related to making final decisions about administrative records.

**Key Assumptions for
Estimated Cost
Savings from
Administrative
Records Are Logical;
the Bureau Will
Continue to Validate
the Assumptions**

As part of our recently released review, we determined that key assumptions in the Bureau's administrative records cost area made sense.¹³ Table 2 shows the results of our analysis.

Table 2: The Census Bureau's Key Assumptions for Estimated Cost Savings from Administrative Records Are Logical and the Bureau Is Taking Steps to Validate Them Further

Key assumption	Support and planned next steps
Removal of vacant and non-residential units will reduce the total Nonresponse Follow-up (NRFU) workload	<p>Percentage reduction based on research examining match rates of records to vacant units in prior census.</p> <p>2015 Census Test successfully demonstrated the ability to remove 11.6 percent of NRFU workload as vacant, corroborating this assumption.</p> <p>Next step: Methods for removing vacant and non-residential units will be tested again in the 2016 Site Test and other tests.</p>
Reduced workload will reduce the total number of field offices needed	<p>Early 2020 planning documents presented a case for reducing the number of field offices.</p> <p>The 2015 Census Test demonstrated reduction in NRFU workload, which drives the number of offices needed.</p> <p>On October 6, 2015, the Bureau announced its decision to replace the 494 field offices in 2010 Census with up to 250 offices in 2020.</p> <p>Next step: The Bureau will test the scale of operations in future tests, including in 2016.</p>
Maximum number of NRFU visits per housing unit will be reduced from 6 to 3 visits	<p>The 2010 Census experiments indicated a reduction in visits might be possible without reducing quality of data.</p> <p>During the test, on multiple occasions, we observed Bureau employees visiting households many more times than the maximum allowed. This calls into question the Bureau's ability to control the number of visits.</p> <p>Next step: Bureau plans to continue testing controls on and impacts of the number of visits during future tests, such as the 2016 Census Test.</p>

¹³GAO-16-48.

Key assumption	Support and planned next steps
Elimination of the Vacant/Delete Check (VDC) operation ^a	<p>The Bureau's 2010 assessment of VDC found the operation had redundancies with NRFU.</p> <p>The 2015 Census Test demonstrated success at identifying vacant and non-residential units, which Bureau officials cite as evidence that this follow-up operation is no longer needed.</p> <p>Next step: Ongoing research to determine the effects of eliminating VDC.</p>
Elimination of the Coverage Follow-up (CFU) operation ^b	<p>Administrative records design for 2020 uses administrative records before and during enumeration, rather than after enumeration, as was done with CFU 2010.</p> <p>Next step: Ongoing research to determine the effects of eliminating the CFU operation.</p>

Source: U.S. Census Bureau documents | GAO-16-206T

^aIn the 2010 Vacant/Delete Check operation, enumerators verified the Census Day status of vacant and deleted (nonexistent) housing units. VDC also attempted to enumerate late additions to the Bureau's address file.

^bDuring 2010 CFU, a contractor telephoned certain households in an attempt to determine if someone had been miscounted.

While we were reviewing these cost assumptions, the Bureau did not always have documentation readily available, and Bureau reporting on one of the assumptions needed to be corrected. We were able to identify the needed support, and Bureau staff said that they will change the methodology for future reporting on the cost estimate to involve more factors and variables, such as the ratio of field workers to supervisors they would need in 2020 in addition to the NRFU workload assumption. This change will help demonstrate the reliability of the estimates as well as ensure effective communication with others about them.

Since we released our report, the Bureau provided an updated estimate of the total 2020 Census life-cycle cost of \$12.5 billion, as well as updated estimates of how much less in four major cost areas its 2020 plan would cost compared to a cost of the a 2020 Census conducted using the 2010 Census approaches and methods. We expect soon to begin reviewing the Bureau's new cost model and its assumption. Bureau officials have told us that although the model has been updated, the key assumptions within the administrative records cost area are largely the same.

Bureau officials told us that the revised life-cycle cost estimate the Bureau released on October 6, 2015 was developed with leading practices from our cost estimating and assessment guide. After the Bureau releases the underlying model, methodology, and supporting documents for the estimate, we anticipate reviewing them to assess their reliability. Continued oversight efforts, such as this hearing, will be helpful to ensure

that the Bureau's efforts remain on track and focused on those most promising to result in a cost-effective 2020 Census.

**GAO Contacts and
Staff
Acknowledgments**

If you have any questions on matters discussed in this statement, please contact Robert Goldenkoff at (202) 512-2757 or by e-mail at goldenkoffr@gao.gov. Other key contributors to this testimony include Ty Mitchell, Assistant Director; Brett Caloia; Robert Gebhart; Richard Hung; Andrea Levine; Donna Miller; Tamara Stenzel; and Timothy Wexler.

Mr. MEADOWS. Thank you.
Ms. Cha.

STATEMENT OF CAROL R. CHA

Ms. CHA. Chairmen Meadows and Hurd, Ranking Members Connolly and Kelly, and members of the subcommittees, I would like to thank you for inviting me to testify today. It's your continued oversight in hearings like these that are vital to ensuring that the 2020 Census is effectively managed.

The Bureau will rely on an enterprise-wide IT initiative called CEDCAP to deliver the systems and infrastructure needed to carry out its redesigned operations. For example, CEDCAP is planning to deliver an online survey instrument and a cloud computing solution to support an Internet response option. For field reengineering, the program is planning to implement a new system to track and manage fieldwork, as well as test the use of mobile devices for field data collection.

Based on our work to date, I'd like to highlight two key challenges regarding the Bureau's plans for 2020. First, the deferral of key IT decisions in the face of time constraints. By August of 2017, the Bureau intends to begin preparations for end-to-end testing to validate that CEDCaP systems are ready to go live on Census Day. This gives the Bureau less than 2 years to develop and integrate planned systems.

While the Bureau has made many key decisions about its redesigned Census as described in its October plan, there are critical IT decisions that have not yet been made, including whether the Bureau will build or buy the needed systems.

This lack of prioritization of IT decisions has been a continuing trend, which we have reported on over the past few years. Most recently, in February we reported that the Bureau had not determined how key research questions would be answered, such as the expected rate of respondents using the Internet survey or the IT infrastructure needed to support this option. As such, we made recommendations to improve the Bureau's ability to answer these questions in time to make those decisions in October. However, this has not happened yet.

Furthermore, the Bureau does not intend to make these decisions and other key ones until 2016 through 2018. Among other things, the build-or-buy decisions won't be made until next June at the earliest. Based on the current sequencing of planned decisions, the Bureau will have about a year to develop and then integrate these systems and then have them ready for end-to-end testing. Further, the mobile device strategy for fieldwork is not expected until 2 months after the start of this testing.

Unless the Bureau makes these key decisions soon, it will likely run out of time to put CEDCaP systems in place. And I can refer you to the screen shot above, which shows examples of deferred 2020 decisions overlaid on top of the Census schedule.

[Slide.]

Ms. CHA. And, again, if you look from left to right, the first one, the build-or-buy decisions, June 2016 at the earliest, and then we come up on Census end-to-end testing, which preparations begin in

August of 2017. So, again, there's very limited time to integrate and implement these systems in time for those tests.

The second challenge is the Bureau's current IT posture. To its credit, important progress has been made to strengthen and institutionalize selected IT management areas, such as governance and requirements management. However, critical IT leadership gaps exist. Most notably, the Bureau is without a permanent chief information officer. Other key vacancies include the chief of the Office Information Security and chief cloud architect. The Bureau is aggressively working to close these gaps. But if they do remain open, its ability to effectively deliver CEDCaP will be hampered.

In addition, the Bureau still has work remaining to fully address our recommendations to improve information security. In January 2013, we made 115 recommendations to address control deficiencies, such as access control to protect its systems from intrusion. As of today, the Bureau has fully addressed 66 of them. The remaining open recommendations, of those, 30 require additional actions by the Bureau and the other 19 are under review. Continued focus on completing this effort must be a high priority to ensure that sensitive information collected during the Census is adequately secure.

In summary, with the deferral of key IT decisions, the Bureau is running out of time to implement the systems needed to support the redesign and achieve its projected \$5.2 billion in cost savings. Moving forward, swift actions to fully implement our open recommendations must be taken. Doing so will improve the Bureau's ability to deliver on its IT plan and realize savings.

That concludes my statement. I look forward to addressing your questions.

[Prepared statement of Ms. Cha follows:]

GAO Highlights

Highlights of GAO-16-205T, a testimony before the Subcommittees on Government Operations and Information Technology, Committee on Oversight and Government Reform, House of Representatives

Why GAO Did This Study

The cost of the nation's decennial census has steadily increased over the past 40 years; the 2010 Census was the most expensive to date, at about \$13 billion. To achieve cost savings while still conducting an accurate count of the population, the U.S. Census Bureau is planning significant changes for the design of the 2020 Decennial Census, including major efforts to implement new technologies and IT systems supporting its surveys. For example, the Bureau is planning to offer an option for households to respond via the Internet, which requires developing new applications and IT infrastructure.

This statement summarizes the critical challenges the Bureau faces in successfully delivering IT systems in time for testing redesigned 2020 Census operations. To develop this statement, GAO relied on previously published work, as well as information on steps the Bureau has taken to implement prior GAO recommendations.

What GAO Recommends

In its prior work, GAO made recommendations to the Census Bureau to prioritize IT testing and research and determine how key decisions for the 2020 Census were to be answered. GAO also made recommendations to improve IT management, workforce planning, and information security. The Bureau has taken steps to address selected recommendations, but more actions are still needed to fully address these recommendations.

View GAO-16-205T. For more information, contact Carol R. Cha at (202) 512-4456 or chac@gao.gov.

November 3, 2015

2020 CENSUS

Key Information Technology Decisions Must be Made Soon

What GAO Found

GAO has previously reported that the U.S. Census Bureau (Bureau) faces a number of critical challenges in developing and deploying the information technology (IT) systems and infrastructure it plans to rely on to conduct the significantly redesigned 2020 Census. Specifically, the Bureau has a major IT program under way to modernize and consolidate the multiple, duplicative systems it currently uses to carry out survey data collection and processing functions; however, with less than 2 years before preparations begin for end-to-end testing of all systems and operations to ensure readiness for the 2020 Census, there is limited time to implement it. While the Bureau documented many key decisions about the redesigned 2020 Census in the 2020 Census Operational Plan, released in October 2015, several key IT-related decisions have not been made. Specifically, the Bureau has not yet made decisions about the projected demand that the IT infrastructure would need to meet or whether it will build or buy the needed systems.

This lack of prioritization of IT decisions has been a continuing trend, which GAO has previously identified. For example:

- In April 2014, GAO reported that the Bureau had not prioritized key IT research and testing needed for its design decisions. Accordingly, GAO recommended that the Bureau prioritize its IT-related research and testing projects. The Bureau had taken steps to address this recommendation, such as releasing a plan in September 2014 that identified research questions intended to inform the operational design decisions.
- In February 2015, however, GAO reported that the Bureau had not determined how key IT research questions that were identified in the September 2014 plan would be answered—such as the expected rate of respondents using its Internet response option or the IT infrastructure that would be needed to support this option. GAO recommended that the Bureau, among other things, develop methodologies and plans for answering key IT-related research questions in time to inform design decisions. However, this has not yet happened.

In addition, while the Bureau has made improvements in some key IT management areas, it still faces challenges in the areas of workforce planning and information security. Specifically:

- It has taken steps to develop an enterprise-wide IT workforce planning process, as GAO recommended in 2012. However, the Bureau has yet to fill key positions. Most concerning, it is currently without a permanent chief information officer.
- The Bureau has taken steps to implement the majority of the 115 recommendations GAO made in 2013 to address information security weaknesses; however, completing this effort is necessary to ensure that sensitive information it will collect during the census is adequately protected.

With the deferral of key IT-related decisions, the Bureau is running out of time to develop, acquire, and implement the systems it will need to deliver the redesign and achieve its projected \$5.2 billion in cost savings.

Chairmen Meadows and Hurd, Ranking Members Connolly and Kelly, and Members of the Subcommittees:

I am pleased to be here today to discuss the U.S. Census Bureau's (Bureau) preparations to implement information technology (IT) solutions for the 2020 Census. On October 6, 2015, the Bureau released the first version of its 2020 Census Operational Plan, which is intended to outline the design decisions that drive how the 2020 Census will be conducted. These design decisions are expected to significantly transform how the Bureau conducts the Decennial Census in an effort to save approximately \$5.2 billion. The redesign largely depends on implementing new technology and systems to modernize and automate many parts of the Decennial Census. Accordingly, concurrent with the overhaul of the 2020 Census, the Bureau is also significantly redesigning the IT systems that support each of its surveys, including the Decennial Census.

With less than 2 years remaining for the Bureau to put all systems and operations place to prepare for end-to-end testing, this hearing is timely. My statement today will describe critical challenges the Bureau faces in successfully delivering the needed IT systems for the 2020 Census.

The information in this testimony is based primarily on our previous reports on the Bureau's planning efforts for the 2020 Census. More detail on our scope and methodology is provided in each published report cited in this testimony. We also obtained and reviewed information on the Bureau's actions in response to our previous recommendations.

The work on which this statement is based was conducted in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

As you know, the cost of the Decennial Census has steadily increased during the past 40 years, in part because the nation's population has steadily grown larger, more diverse, and increasingly difficult to enumerate. For example, at about \$13 billion, the 2010 Census was the costliest U.S. census in history and was 56 percent more costly than the \$8.1 billion 2000 Census (in constant 2010 dollars).

To help save costs, in preparing for the 2020 Census, the Bureau has been researching and testing new methods and technologies to redesign the Census to more cost-effectively count the population while maintaining high-quality results.

The Bureau's research and testing has focused on four redesign areas:

- **Reengineering address canvassing:** This involves reengineering processes for updating the Bureau's address list and maps of the nation to reduce the need for employing field staff to walk every street in the nation to verify addresses.
- **Optimizing self-response:** Includes efforts to maximize the self-response of households by, among other things, offering an Internet response option. As we have previously reported, to deliver the Internet response option, the Bureau would need to, among other things, design and develop an Internet response application, develop and acquire the IT infrastructure to support a large volume of data processing and storage, and plan communication and outreach strategies to motivate households to respond via the Internet.¹
- **Using administrative records:** This includes expanding the use of data previously obtained by other federal and state government agencies and commercial sources to reduce the need for costly and labor-intensive follow-up work. My colleague will address the Bureau's progress on using administrative records in his statement, today.²
- **Reengineering field operations:** This includes reducing the number of visits to households, automating the management of enumerator work to conduct non-response follow-up, and automating and optimizing case assignment and routing for enumerators to reduce the staffing, infrastructure, and field offices required for the 2020 Census.

The Bureau has conducted several major field tests to examine the potential for each of these redesign areas:

- In mid-2014 the Bureau conducted the 2014 Census Test in the Maryland and Washington, D.C., areas to test new methods for conducting self-response and non-response follow-up.

¹GAO, *2020 Census: Key Challenges Need to Be Addressed to Successfully Enable Internet Response*, GAO-15-225 (Washington, D.C.: Feb. 5, 2015).

²GAO, *2020 Census: Progress Report on Using Administrative Records to Control Enumeration Costs*, GAO-16-206T (Washington, D.C.: Nov. 3, 2015).

-
- In early 2015 the Bureau completed the Address Validation Test, which was used to examine new methods for updating the Bureau's address list.
 - In mid-2015 the Bureau conducted the 2015 Census Test in Arizona to test, among other things, the use of a field operations management system to automate data collection operations and provide real-time data and the ability to reduce the non-response follow-up workload using data previously provided to the government, as well as enabling enumerators to use their personally owned mobile devices to collect census data.
 - Also in mid-2015, the Bureau conducted an optimizing self-response test in Savannah, Georgia, and the surrounding area, which was intended to further explore methods of encouraging households to respond using the Internet, such as using advertising and outreach to motivate respondents, and enabling households to respond without a Bureau-issued identification number.
 - More recently, the Bureau began its National Content Test, which is currently ongoing and intended to, among other things, continue to test self-response modes and contact strategies and refine estimates of national self-response and Internet response rates.

These tests were intended to inform the first version of the Bureau's 2020 Census Operational Plan, which is intended to outline design decisions that drive how the 2020 Census will be conducted. As part of these decisions, the Bureau has committed to aspects of the 2020 Census redesign. The operational plan articulated 326 total design decision points, which vary widely in their complexity, importance, and urgency.³ As of October 6, 2015, the Bureau had made decisions for about 47 percent of them related to each of the four redesign areas. For example, the Bureau has decided to

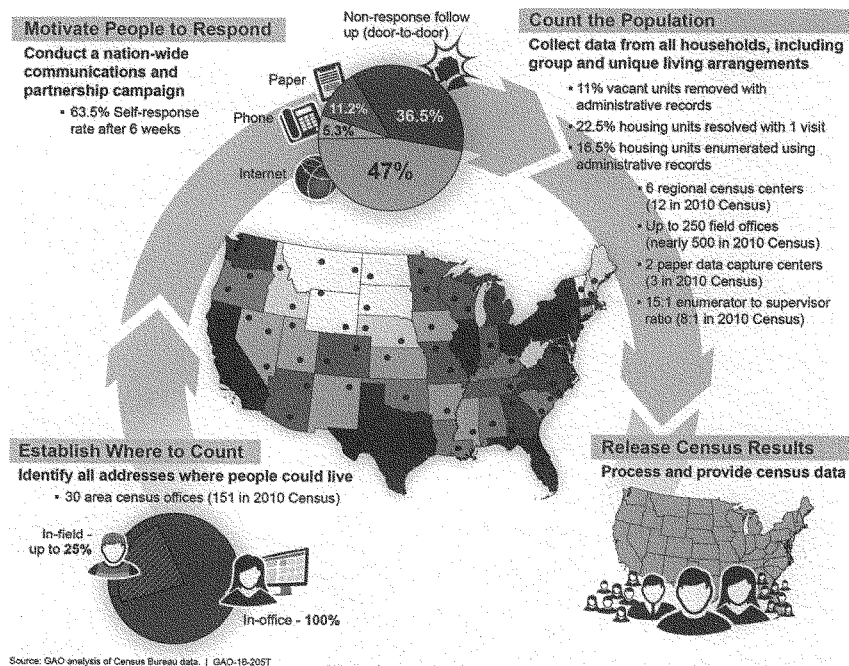
- conduct 100 percent of address canvassing (i.e., identifying all addresses where people could live) in the office, and target a subset of up to 25 percent for in-the-field address canvassing;
- offer an Internet self-response option, as well as alternative response options via telephone and paper for limited circumstances;

³This excludes 25 decisions in the Program Management operational area because they are related to management of the program rather than design elements of the 2020 Census operations.

-
- allow people to respond without a unique census identification number;
 - use mobile devices for enumerators to conduct field data collection;
 - use administrative records to enumerate vacant units;
 - use enterprise solutions to support the 2020 Census, when practicable; and
 - reduce the field footprint by half in comparison to the 2010 Census (e.g., 6 regional census centers instead of 12 and up to 250 field offices instead of nearly 500).

Figure 1 provides an overview of the Bureau's current plans and assumptions for the 2020 Census, resulting from the October 2015 operational plan.

Figure 1: Overview of the Census Bureau's Plans and Assumptions for the 2020 Census, as of October 6, 2015



As a result of these decisions, the Bureau estimates saving \$5.2 billion. Specifically, the Bureau estimated that repeating the design of the 2010 Census for 2020 would cost approximately \$17.8 billion (in constant 2020 dollars), while successfully implementing the four redesign areas is expected to result in an overall 2020 Census cost of \$12.5 billion (in

constant 2020 dollars). Table 1 illustrates the estimated cost savings associated with each redesign area.

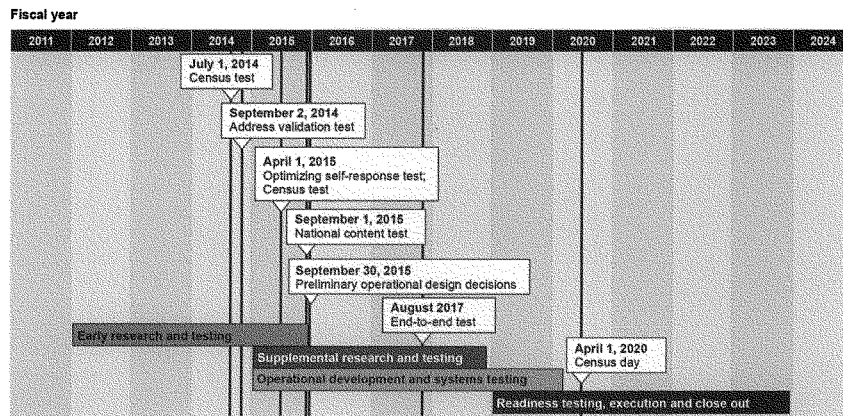
Table 1: 2020 Census Redesign Areas and Estimated Savings as of October 6, 2015

Initiative	Bureau's estimated cost savings (in billions)
Reengineering address canvassing	\$0.9
Optimizing self-response	\$0.4
Utilizing administrative records	\$1.4
Reengineering field operations	\$2.5
Total	\$5.2

Source: Census Bureau | GAO-16-205T

Moving forward, the Bureau plans to conduct additional research and testing and further refine the design through 2018. By August 2017, the Bureau plans to begin preparations for end-to-end testing, which is intended to test all systems and operations to ensure readiness for the 2020 Census. Figure 2 shows the timeline for planned 2020 Census research and testing.

Figure 2: 2020 Decennial Census Planned Research and Testing Schedule



Source: GAO analysis of Census Bureau data. | GAO-16-205T

Note: The Census Bureau had planned to complete the preliminary operational design decisions by September 30, 2015, but they were released on October 6, 2015.

Concurrent with redesigning the decennial census, the Bureau has also begun a significant effort to modernize and consolidate its survey data collection and processing functions. This is being undertaken through an enterprise-wide IT initiative called Census Enterprise Data Collection and Processing (CEDCAP). This initiative is a large and complex modernization program intended to deliver a system-of-systems for all the Bureau's survey data collection and processing functions—rather than continuing to rely on unique, survey-specific systems with redundant capabilities.

For the 2020 Census, CEDCAP is expected to deliver the systems and IT infrastructure needed to implement the Bureau's redesign areas. For example:

- To reengineer field work, CEDCAP is expected to implement a new dynamic operational control system to track and manage field work. This system is to be able to make decisions about which visits enumerators should attempt on a daily basis using real-time data, as

well as provide automated route planning to make enumerator travel more efficient. CEDCAP also includes testing the use of mobile devices, either government-furnished or employee-owned, to automate data collection in the field.

- To maximize self-response with the Internet response option, CEDCAP is responsible for developing and testing a web-based survey application and exploring options for establishing the IT infrastructure to support the increased volume of data processing and storage that will be needed.

CEDCAP consists of 12 projects that are to deliver capabilities incrementally, over the course of at least 10 releases. The Bureau plans to roll out capabilities for the 2020 Census incrementally through 6 of these releases, while also deploying capabilities for other surveys such as the American Community Survey and Economic Census. The Bureau expects to reuse selected systems, make modifications to other systems, and develop or acquire additional systems and infrastructure. As of August 2015, the CEDCAP program was projected to cost about \$548 million through 2020.

However, the Bureau's past efforts to implement new approaches and systems have not always gone well. As one example, during the 2010 Census, the Bureau planned to use handheld mobile devices to support field data collection for the census, including following up with nonrespondents. However, due to significant problems identified during testing of the devices, cost overruns, and schedule slippages, the Bureau decided not to use the handheld devices for non-response follow-up and reverted to paper-based processing, which increased the cost of the 2010 Census by up to \$3 billion and significantly increased its risk as it had to switch its operations to paper-based operations as a backup.⁴

Bureau Faces Critical Challenges in Delivering IT Systems Needed to Support Redesign Areas

Last month's issuance of the 2020 Census Operational Plan, which documents many key decisions about the redesign of the 2020 Census, represents progress; however, the Bureau faces critical challenges in

⁴GAO, *2010 Census: Preliminary Lessons Learned Highlight the Need for Fundamental Reforms*, GAO-11-496T (Washington, D.C.: Apr. 6, 2011); and *Information Technology: Census Bureau Testing of 2010 Decennial Systems Can Be Strengthened*, GAO-09-262 (Washington, D.C.: Mar. 5, 2009).

delivering the IT systems needed to support the redesign areas. Specifically, with preparations for end-to-end testing less than 2 years away, the window to implement CEDCAP, which is intended to be the backbone of the 2020 Census, is narrow. Additionally, while the Bureau has demonstrated improvements in IT management, as we have previously reported, it faces critical gaps in its IT workforce planning and information security. Until it takes actions we have previously recommended to address these challenges,⁵ the Bureau is at risk of cost overruns, schedule delays, and performance shortfalls, which will likely diminish the potentially significant cost savings that it estimates will result from redesigning the census for 2020.

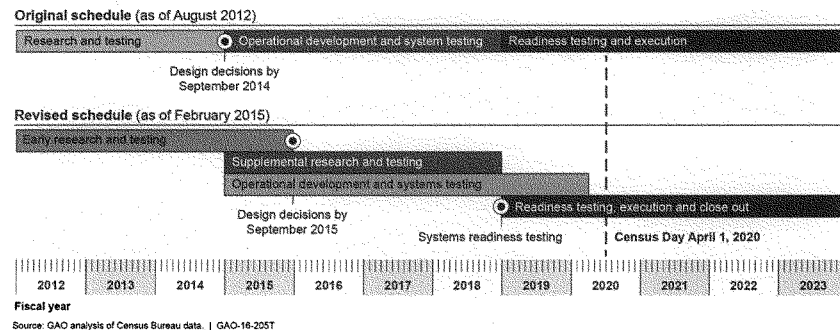
Timeframe to Implement the Large and Complex IT Systems Needed for 2020 Census Redesign Is Narrow

The Bureau has not prioritized key IT-related decisions, which is a trend we have reported for the past few years. Specifically, in April 2014, we reported the Bureau had not prioritized key IT research and testing needed for the design decisions planned for the end of 2015.⁶ In particular, the Bureau had not completed the necessary plans and schedules for research and testing efforts and had not prioritized what needed to be done in time for the 2015 design decisions—a milestone that had already been pushed back by a year (see fig. 3). We concluded that, given the current trajectory and the lack of supporting schedules and plans, it was unlikely that all planned IT-related research and testing activities would be completed in time to support the 2015 design decisions—which ultimately came to fruition (discussed later).

⁵GAO, *Information Technology: Census Bureau Needs to Implement Key Management Practices*, GAO-12-915 (Washington, D.C.: Sept. 18, 2012); and *Information Security: Actions Needed by Census Bureau to Address Weaknesses*, GAO-13-63 (Washington, D.C.: Jan. 22, 2013). Another version of this report was issued for limited distribution. GAO, *2020 Census: Prioritized Information Technology Research and Testing Is Needed for Census Design Decisions*, GAO-14-389 (Washington, D.C.: Apr. 3, 2014); GAO-15-225.

⁶GAO-14-389.

Figure 3: Original and Revised Schedules for 2020 Decennial Census Design Decision



Source: GAO analysis of Census Bureau data. | GAO-16-205T

Note: The Census Bureau had planned to complete design decisions by September 2015, but the decisions were released on October 6, 2015.

In light of these ongoing challenges, we recommended in our April 2014 report that the Bureau prioritize its IT-related research and testing projects that need to be completed to support the design decisions and develop schedules and plans to reflect the new prioritized approach. The Bureau agreed with our recommendations and has taken steps to address them. For example, in September 2014, the Bureau released a plan that identified inputs, such as research questions, design components, and testing, that were needed to inform the operational design decisions expected in the fall of 2015.

However, as we reported in February 2015,⁷ the Bureau had not yet determined how key IT research questions that had been identified as critical inputs into the design decisions—estimating the Internet self-response rate and determining the IT infrastructure for security and scalability needed to support Internet response—were to be answered. We therefore recommended that the Bureau, among other things, develop methodologies and plans for answering key IT-related research questions in time to inform key design decisions.

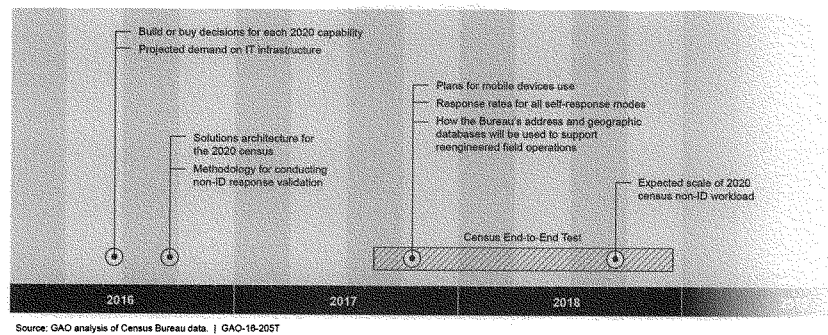
⁷GAO-15-225.

While the recent 2020 Census Operational Plan documents many key IT-related decisions about the redesign of the census, other critical questions, including the ones identified in our February 2015 report, remain unanswered. Of greater concern, the Bureau does not intend to answer these and other questions until 2016 through 2018. Specifically, there are several significant IT decisions that are being deferred, which have implications on the CEDCAP program's ability to have production-ready systems in place in time to conduct end-to-end testing. For example, the Bureau does not plan to decide on

- the projected demand that the IT infrastructure and systems would need to accommodate or whether the Bureau will build or buy the needed systems until June 2016, at the earliest;
- the high-level design and description of the systems (referred to as the solutions architecture) until September 2016—leaving about a year to, among other things, build or acquire, integrate, and test the systems that are intended to serve as the backbone to the 2020 Census before preparations for end-to-end testing begins in August 2017; and
- the strategy for the use of mobile devices for field work until October 2017.

Figure 4 illustrates several key IT-related decisions that have been deferred which will impact preparations for the end-to-end test and 2020 Census. Unless the Bureau makes these key decisions soon, it will likely run out of time to put CEDCAP systems in place.

Figure 4: Examples of Deferred 2020 Census-related IT Decisions Impacting CEDCAP



Census Bureau Has Demonstrated Improvements in IT Management, but Faces Critical Gaps in IT Workforce and Information Security

Institutionalizing key IT management controls, such as IT governance, system development methodology, and requirements management processes, helps establish a consistent and repeatable process for managing and overseeing IT investments and reduces the risk of experiencing cost overruns, schedule slippages, and performance shortfalls, like those that affected the previous census.

However, in September 2012,⁸ we reported that the Bureau lacked a sufficiently mature IT governance process to ensure that investments are properly controlled and monitored, did not have a comprehensive system development methodology, and continued to have long-standing challenges in requirements management. We made several recommendations to address these issues, and the Bureau took actions to fully implement each of the recommendations. For example, the Bureau

- addressed gaps in policies and procedures related to IT governance, such as establishing guidelines on the frequency of investment review

⁸GAO-12-915.

board meetings and thresholds for escalation of cost, risk, or impact issues;

- finalized its adoption of an enterprise system development life-cycle methodology, which included the short incremental development model, referred to as Agile, and a process for continuously improving the methodology based on lessons learned; and
- implemented a consistent requirements development tool that includes guidance for developing requirements at the strategic mission, business, and project levels and is integrated with its enterprise system development life-cycle methodology.

As a result, the Bureau has established a consistent process for managing and overseeing its IT investments.

IT Workforce Planning Has Improved, but Critical Gaps Continue to Exist

Effective workforce planning is essential to ensure organizations have the proper skills, abilities, and capacity for effective management. While the Bureau has made progress in IT workforce planning efforts, many critical IT competency gaps remain to be filled. In September 2012 we reported, among other things, that the Bureau had not developed a Bureau-wide IT workforce plan; identified gaps in mission-critical IT occupations, skills, and competencies; or developed strategies to address gaps.⁹

Accordingly, we recommended that the Bureau establish a repeatable process for performing IT skills assessments and gap analyses and establish a process for directorates to coordinate on IT workforce planning. In response, in 2013 the Bureau completed an enterprise-wide competency assessment and identified several mission-critical gaps in technical competencies. In 2014, the Bureau established documents to institutionalize a strategic workforce planning process, identified actions and targets to close the competency gaps by December 2015, and established a process to monitor quarterly status reports on the implementation of these actions.

However, as we reported in February 2015, while these are positive steps in establishing strategic workforce planning capabilities, the Bureau's workforce competency assessment identified several mission-critical gaps that would challenge its ability to deliver IT-related initiatives, such as the IT systems that are expected to be delivered by CEDCAP.¹⁰ For example,

⁹GAO-12-915.

¹⁰GAO-15-225.

the Bureau found that competency gaps existed in cloud computing, security integration and engineering, enterprise/mission engineering life-cycle, requirements development, and Internet data collection. The Bureau also found that enterprise-level competency gaps existed in program and project management, budget and cost estimation, systems development, data analytics, and shared services.

The Bureau has taken steps to regularly monitor and report on the status of its efforts to close competency gaps and has completed several notable actions. For example, in August 2015, the Bureau filled the position of Decennial IT Division Chief and in September 2015, awarded an enterprise-wide IT services contract for systems engineering and integration support.

However, more work remains for the Bureau to close competency gaps critical to the implementation of its IT efforts. Most significantly, in July 2015, the Chief Information Officer resigned. As of October 2015, the Bureau was working to fill that gap and had an acting Chief Information Officer temporarily in the position.

Additionally, there are other gaps in key positions, such as the Chief of the Office of Information Security and Deputy Chief Information Security Officer, Big Data Center Chief, and Chief Cloud Architect, and the CEDCAP Assistant Chief of Business Integration, who is responsible for overseeing the integration of schedule, risks, and budget across the 12 projects. According to Bureau officials, they are working to address these gaps.

More Work Remains to Address Information Security Weaknesses

Critical to the Bureau's ability to perform its data collection and analysis duties are its information systems and the protection of the information they contain. A data breach could result in the public's loss of confidence in the Bureau, thus affecting its ability to collect census data. To ensure the reliability of their computerized information, agencies should design and implement controls to prevent, limit, and detect unauthorized access to computing resources, programs, information, and facilities.¹¹ Inadequate design or implementation of access controls increases the risk of unauthorized disclosure, modification, and destruction of sensitive information and disruption of service.

¹¹Access controls include those related to (1) protection of system boundaries, (2) identification and authentication, (3) authorization, (4) cryptography, (5) audit and monitoring, and (6) physical security.

In January 2013, we reported on the Bureau's implementation of information security controls to protect the confidentiality, integrity, and availability of the information and systems that support its mission.¹² We concluded that the Bureau had a number of weaknesses in controls intended to limit access to its systems and information, as well as those related to managing system configurations and unplanned events. We attributed these weaknesses to the fact that the Bureau had not fully implemented a comprehensive information security program, and made 115 recommendations aimed at addressing these deficiencies.¹³ The Bureau expressed broad agreement with the report and said it would work to find the best ways to address our recommendations.

As of October 29, 2015, the Bureau had addressed 66 of the 115 recommendations we made in January 2013. Of the remaining open recommendations, we have determined that 30 require additional actions by the Bureau, and for the other 19 we have work under way to evaluate if they have been fully addressed. The Bureau's progress toward addressing our security recommendations is encouraging. However, more work remains to address the recommendations.

A cyber incident recently occurred at the Bureau, and while it appears to have had limited impact, it demonstrates vulnerabilities at the Bureau. Specifically, in July 2015, the Bureau reported that it had been targeted by a cyber attack aimed at gaining access to its Federal Audit Clearinghouse, which contains non-confidential information from state and local governments, nonprofit organizations, and Indian tribes to facilitate oversight of federal grant awards. According to Bureau officials, the breach was limited to this database on a segmented portion of the Bureau's network that does not touch administrative records or sensitive respondent data protected under Title 13 of the U.S. Code, and the hackers did not obtain the personally identifiable information of census and survey respondents.

Given that the Bureau is planning to build or acquire IT systems to collect the public's personal information for the 2020 Census in ways that it has not for previous censuses (e.g., web-based surveys, cloud computing, and enabling mobile devices to collect census data), continuing to

¹²GAO-13-63. Another version of this report was issued for limited distribution.

¹³This included 13 public recommendations and 102 recommendations for limited distribution.

implement our recommendations and apply IT security best practices as it implements CEDCAP systems must be a high priority.

As a result of the Bureau's challenges in key IT internal controls and looming deadline, we identified CEDCAP as an IT investment in need of attention in our February 2015 High-Risk report.¹⁴ We recently initiated a review of the CEDCAP program for your subcommittees, and expect to issue a report in the spring of 2016.

In conclusion, the Bureau is pursuing initiatives to significantly reform its outdated and inefficient methods of conducting decennial censuses. However, with less than 2 years remaining until the Bureau plans to have all systems and processes for the 2020 Census developed and ready for end-to-end testing, it faces challenges that pose significant risk to 2020 Census program. These include the magnitude of the planned changes to the design of the census, the Bureau's prior track record in executing large-scale IT projects, and the current lack of a permanent Chief Information Officer, among others. Moreover, the Bureau's preliminary decision deadline has come and gone, and many IT-related decisions have been deferred to 2016 through 2018. Consequently, it is running out of time to develop, acquire, and implement the production systems it will need to deliver the redesign and achieve its projected \$5.2 billion in cost savings.

The Bureau needs to take action to address the specific challenges we have highlighted in prior reports. If these actions are not taken, cost overruns, schedules delays, and performance shortfalls may diminish the potentially significant cost savings that the Bureau estimates will result from redesigning the census for 2020.

Chairmen Meadows and Hurd, Ranking Members Connolly and Kelly, and Members of the Subcommittees, this completes my prepared statement. I would be pleased to respond to any questions that you may have.

¹⁴Every 2 years at the start of a new Congress, GAO calls attention to agencies and program areas that are high risk due to their vulnerabilities to fraud, waste, abuse, and mismanagement, or are most in need of transformation. As part of a new entry into the February 2015 update to our High-Risk Series focused on improving the management of IT acquisitions and operations, CEDCAP was identified as an example of an IT investment—among others across the federal government—in need of the most attention. See GAO, *High-Risk Series: An Update*, GAO-15-290 (Washington, D.C.: Feb. 11, 2015).

Mr. MEADOWS. Thank you so much.

The chair recognizes the chairman of the Subcommittee on Information Technology, Mr. Hurd, for 5 minutes for a series of questions.

Mr. HURD. Thank you, Chairman.

And thank you all for your testimony today.

And, Ms. Cha, I want to thank you for your written statement. It was very good in helping to explain some of the unique challenges that we're having to provide oversight on.

My first question, I guess, is first to you, Director Thompson. We've talked about the lack of a permanent CIO. We talked about the lack of a chief of the business integration, which is managing the CEDCaP program. What are the plans to fill those spots? What are the challenges that you're facing in getting people there? And who is filling that role as a temporary basis?

Mr. THOMPSON. Thank you, Congressman.

So the job announcement for the CIO is out and available now. CIO Cooper and I have been widely circulating it so that we can be sure that we get a good number of applicants. Harry Lee, sitting behind me, has done an excellent job in the interim. He's the deputy CIO. And he's been doing a great job filling the vacancy in the absence, and we've made a number of accomplishments in the IT area.

Mr. HURD. How long have these spots been open?

Mr. THOMPSON. It became open in July.

Mr. HURD. Okay. And we have applicants putting information in or submitting resumes or whatever the process is?

Mr. THOMPSON. So I don't know that I can talk about the applicants at this point in the process.

Mr. HURD. Mr. Cooper, did you have something to add?

Mr. COOPER. Yes. With the position open, it will remain open for 30 days. Our human resources group will do the initial processing of those applicants and those who are well qualified will be passed to the appropriate review panel. We'll then make that ultimate decision about selecting a CIO.

Mr. HURD. Because my concern is—

Mr. COOPER. So we have not yet seen any of the responses. I don't have any information on that.

Mr. HURD. Thank you. My concern is with the lack of an identified leader in managing this process, if CEDCaP is going to be the process, and we have less than 2 years, we need to make sure that that person is in place.

My next set of questions is there was 115 recommendations, as Ms. Cha pointed out in her written and verbal testimony. We have information that 19 of those 115 have been addressed and work has started on 66, or 66 of those have been addressed. Can you give me an update?

Mr. THOMPSON. Certainly. So last April, when we received—well, last April we responded to all 115 recommendations that the GAO made, and we've been working with the GAO since that point. To date, we believe that they have closed 65 of them. We believe that they referred 19 back to us. And based on our discussions, we're still waiting to hear from them on about 30 more. So that's the current status from our records.

Mr. HURD. Ms. Cha, do you have any insights or comments on that?

Ms. CHA. Sure. So as of today, 66 have been closed out. But I believe Director Thompson has those numbers transposed. There are 19 that are currently under review and there are 30 remaining recommendations that still require additional actions by the Bureau.

Mr. HURD. What's the biggest priority one of those 30 that require action?

Ms. CHA. Sure. So in terms of the open recommendations, they relate to the identification and authentication types of controls. So that would be like password controls, unsecured system accounts and access, as well as configuration management. So that would be like——

Mr. HURD. All the things that have led to a number of breaches not only in the private sector, but in the Federal Government.

Ms. CHA. That's correct.

Mr. HURD. And I'm going to address this to you, Mr. Cooper, since this falls in your bailiwick. What's being done to address that, on the authentication and the patch management programs?

Mr. COOPER. Okay. We've taken a couple steps proactively following the recent OMB 30-day cyber——

Mr. HURD. Let me rephrase the question. When is this going to be completed?

Mr. COOPER. I wish I could give you an exact date. We have initiated a whole series of activities that are targeted for completion by 31 December of this fiscal year—I'm sorry, this calendar year. That does include the open actions from the GAO report.

We've also taken an additional step, and that is that we've created internal cross-bureau teams of our cyber experts, and we are bringing those to bear to assist all of our bureaus but, in this case the Census Bureau, in helping close these actions. So in addition to the Census Bureau's cyber experts, we've brought additional expertise from within the Department. We believe that that will help us both meet our target deadline and bring additional skills and expertise objectively to bear on these open items.

Mr. HURD. Thank you.

And my last question to you, Director Thompson, and I'm running out of time, I've run out of time, if you realize savings from using new technology or a different process, are you able to use that savings within the calendar or the fiscal year on other projects?

Mr. THOMPSON. Congressman——

Mr. HURD. And that's not a trick question.

Mr. THOMPSON. No, no, I know——

Mr. HURD. Because I'm trying to figure out, if you're doing things, you're realizing savings, I want you to be able to use those savings to go back in to do your work.

Mr. THOMPSON. So we are not going to be asking for anything more than the money we need to do the Census, and we believe that we are going to avoid \$5.2 billion in costs. In saying that, we also plan to make sure that we have the right resources in the right areas to make sure we count everyone as well.

Mr. HURD. Mr. Chairman, I apologize, I yield back the time that I do not have.

Mr. MEADOWS. No, I thank the gentleman for his insightful questions.

The chair recognizes the gentleman from the 11th District of Virginia, Mr. Connolly.

Mr. CONNOLLY. Thank you, Mr. Chairman. And I want to point out I think three of our witnesses are Virginians and all three confirmed under oath that they voted.

Now, Mr. Thompson, where are you from.

Mr. THOMPSON. I'm from Old Town Alexandria.

Mr. CONNOLLY. All right. So you're a Virginian. All right. All four. Bless them. Wonderful people doing wonderful work, let's just be honest.

Anyway, thank you, Mr. Chairman.

First of all, I want to say to all of you on the panel, you've got four people up here who actually care about the mission and are focused on it and not allowing a lot of extraneous issues to get in our way of trying to be supportive in a rational way.

And I would just note that, because I really appreciate my friends, Ms. Kelly and Mr. Meadows and Mr. Hurd, for really being in a collaborative spirit, and we're collaborating on especially a lot of other IT issues too. I just want to say that because it doesn't happen as often as we'd like up here and it hardly ever gets recognized when it does. And so I hope we can take this as a magic moment and try to work together with the executive branch on the Census.

Yeah, that's right.

Director Thompson, help us understand, how much money do you believe the Census requires to undertake its mission this next time? We're trying to follow the numbers.

Mr. THOMPSON. So we estimate that we're going to, for the full cycle cost of the Census' interoperational plan, we're going to require \$12.3 billion.

Mr. CONNOLLY. \$12.3. And is that what you requested in the budget?

Mr. THOMPSON. Well, we request the budget on a yearly basis.

Mr. CONNOLLY. I know.

Mr. THOMPSON. We are moving towards that with our request for fiscal year 2016.

Mr. CONNOLLY. And what do you project right now the shortfall between the request—I think I cited \$375 million, but that's probably just for 1 year—what is the shortfall between what you need and so far what has been appropriated or what you project to be appropriated?

Mr. THOMPSON. Congressman, I think you hit the number on the head when you mentioned the shortfall that we initially saw for fiscal year 2016. However, we haven't seen the final appropriation—

Mr. CONNOLLY. We haven't either.

Mr. THOMPSON. —for fiscal year 2016.

Mr. CONNOLLY. We haven't either.

Are you able to disaggregate for the purposes of funding the Census what you think you would get appropriated pursuant to the

new budget agreement we passed last week? We increased domestic discretionary investments by roughly \$25 billion for this next fiscal year in that budget agreement, 25 for defense, 25 for domestic. Any idea how much of your projected shortfall, in an ideal world, I know it's hard because we haven't appropriated it yet, but any sense of whether that would help the funding shortfall?

Mr. THOMPSON. Congressman, the only numbers I'm familiar with are the President's budget and the actions that have taken place.

Mr. CONNOLLY. Okay. Well, I think it's going to be really important to follow that and have better answers. I can't ask you to have a better answer yet. We don't either. But if the four of us are going to be helpful and supportive—

Mr. THOMPSON. Right. But we are preparing to work with the Congress to explain exactly how we'll spend every dollar that we get—

Mr. CONNOLLY. There was a method to my madness in telling you you've got four people here who actually want to get something done and be productive. And part of that is the budget. But we got to know—we've all got to be on the same page in terms of those projected numbers.

Mr. Cooper, thank you for your leadership as CIO. Would that we could replicate you in lots of other Federal agencies. Help us have confidence. In 2010, we had a handheld device meltdown. We're doubling down on those devices for this next go-round. What makes us confident that we're, in fact, going to make it work and save money, as opposed to another meltdown, only double?

Mr. COOPER. Okay. The first observation would be that in 2010—and I was not part of the Department in 2010, so the information I'm sharing is information I've gained since coming onboard. If I don't get something exactly right or if you need more detail, I'm more than happy to come back up, set up briefings, fill in the details. But based upon what I've investigated and what I've kind of been involved and learned, the first important difference, the technology that was used or attempted to be used in 2010 was very specific, specialized, proprietary technology.

Now, compare and contrast that. In 2020, we're using commercially available solutions. They're off the shelf. They've been proven in industry. We're talking about smartphones, old devices. I think probably every one of us in this room carry some type of phone, smartphone type of device. We may carry a tablet, that type of thing. So a huge difference. This is not unproven, single-use, proprietary technology.

What also comes with that? Industry itself, all of the carriers, major telecom carriers, as well as the device providers, have invested significant amounts of expertise, research and development dollars into helping secure those devices.

So, again, this is not something where the Census Bureau by itself takes the lead responsibility along with its industry partners. That's a huge, important, significant difference in my ability to say to you with confidence we're not repeating a similar type of scenario. We're using proven technology, well understood, well proven in the marketplace. We are then, additionally, applying our own NIST-driven security protocols, risk framework, cybersecurity

framework, GAO recommendations, IG recommendations all come into play, so that we have a significantly higher degree of confidence in being able to use those platforms in field operations.

Mr. CONNOLLY. Thank you very much.

Thank you, Mr. Chairman.

Mr. MEADOWS. The chair recognizes the gentleman from Iowa, Mr. Blum, for 5 minutes of questioning.

Mr. BLUM. Thank you, Chairman Meadows.

Thank you to the panel today for sharing your insights with us.

They say that genius is the ability to reduce the complicated to the simple. And it would be great to have a little genius in the Federal Government. So toward that end, Mr. Cooper, do we have a database of all the addresses in the United States of America.

Mr. COOPER. My understanding is the closest we have to that in the Census Bureau is the Master Address File. I believe that in addition to that there are supplemental databases, also administrative records that the Director has talked about in some of the earlier GAO testimony. Therefore, I'm going to say, but I'm going to ask for help from my fellow panelists, I don't believe that there's a single master file of every housing unit in the United States.

But is that in the Master Address File?

Mr. BLUM. Would the Postal Service have that?

Mr. THOMPSON. Can I respond?

Mr. BLUM. Absolutely.

Mr. THOMPSON. We maintain a master address list that we try to make as accurate as possible, that contains every address in the United States, both mailable addresses and those that aren't mailable. And we are in the process now of continuously updating it. So we work with the States and localities that want to participate and take their lists and put into it. Twice a year at least we get a feed from the post office and they give us updates to their delivery sequence file. So we try to maintain this address list and make it as accurate as possible so that when we take the Census we have every housing unit in there.

Mr. BLUM. Safe to say we have 99 percent of the addresses in the country in the Postal System's database? Ninety-eight percent? High percentage?

Mr. THOMPSON. So we actually believe that we have an address list that is a little more complete than the post office's list because we actually go into structures that don't receive mail and list the addresses. And part of what we're going to do before the next Census in about 25 percent of the country is actually go knock on doors, ask if there are hidden addresses, to make sure that we get every address in the country where someone can reside, whether or not they get mail at that address.

Mr. BLUM. And your job is to count the number of people living at each address, correct, simply put?

Mr. THOMPSON. Yes.

Mr. BLUM. I just want to make sure. Maybe I'll ask Mr. Cooper this. Do you think we're unnecessarily complicating this? I come from the high-tech industry, the private sector. I know how this can go.

Mr. COOPER. My direct answer would be no, sir, I don't believe that we are overly complicating it. But here is what's going on.

We're moving from pencil-and-paper processing to leveraging technology in a more fully automated way. So the transition is complicated, complex, and we need to manage that extremely well and extremely rigorously and thoroughly. But I do not believe that we are making it more complicated than it needs to be.

Mr. BLUM. Mr. Cooper, I think it's called CEDCaP?

Mr. COOPER. Yes, sir.

Mr. BLUM. I read there was a mention of the last Census, I believe we had over 100 duplicative systems? To me, that's, A, amazing, and, B, a recipe for disaster. How are we changing that this Census?

Mr. COOPER. CEDCaP, I admit, Congressman, I don't know the exact numbers. So I accept the numbers that you're saying on what did exist and the number of duplicative systems.

Here's the really good news. CEDCaP, in fact, is being architected, engineered, and it is well on the way to replace all of those, at least as many as are appropriate to the 2020 decennial and the concomitant operations, to replace that with a single architected, well understood, not complicated environment ecosystem, if you will, that will serve all data collection and all operational processing of that data in the conduct of the 2020 decennial.

Mr. BLUM. The last thing, my time is running short, as I was reading, it seems to me the largest savings is in the Nonresponse Followup organization, \$2.5 billion in savings, which would be fantastic. The wording in there said that we've developed or are developing a prototype system that incorporates commercial off-the-shelf technology, which I think is great, rather than developing it custom. Can you talk to me about that?

Mr. COOPER. As much as I have been fully briefed and I understand what I've been told, which are two different things, the fact that we are using commercial off-the-shelf solutions is certainly something that I, in my CIO role, are fully supportive of. Based on your reflection on your background, I fully concur with you.

By using commercially proven off-the-shelf software, same thing that I testified just briefly about, the mobile devices, mobile technology, that type of thing, those same benefits apply, okay? We have proven technology. We've got a lot of additional expertise that has been brought to bear. It usually incorporates architected security features that we then can take advantage of. They're well known, well proven across the industry.

Wherever we can, and it's in the operational plan that was recently shared with Congress and released publicly, we are leveraging commercial off-the-shelf software and solutions.

Mr. BLUM. Excellent.

And I'll yield back the time I do not have, Chairman Meadows.

Mr. MEADOWS. I thank the gentleman.

The chair recognizes the gentlewoman from Illinois, Ms. Kelly, for 5 minutes.

Ms. KELLY. Thank you, Mr. Chair.

The CEDCaP program is intended to help the cost-saving efforts for 2020. Mr. Cooper, your written testimony for today notes that CEDCaP, "aims to create an integrated and standardized system-of-systems that will offer shared data collection and processing

across all Census Bureau operations,” and you seem very positive about the program.

Ms. Cha, what steps should the Bureau take to ensure that CEDCaP systems are ready for planned end-to-end testing.

Ms. CHA. Sure. So with CEDCaP, the system-of-systems approach, there are 12 projects underneath it. And what we find across the Federal Government with these troubled or failed IT projects, the common thread behind all of these is that they’re simply too large in scope. And so, obviously, with FITARA and the best practices that GAO endorses, going incremental is the right approach. You want to chunk out these large scope efforts into these smaller pieces that are more manageable in nature.

So in looking at the scope of CEDCaP itself, it’s simply too large given the time remaining. And the concern that we have from GAO’s perspective is we’re getting to a point where time is going to drive the technology that can be delivered for 2020, as opposed to business requirements. So the absolute first thing that the Bureau should be doing is to make these key IT decisions as soon as possible.

Ms. KELLY. And what would be the effect if the Bureau does not have that CEDCaP systems in place before end-to-end testing begins?

Ms. CHA. Right. Well, the end-to-end testing is absolutely critical ensuring that—in terms of an effect, you know, most recently the most high profile effect that we have seen is healthcare.gov, where they gave short shrift to that end-to-end testing. And so if the schedules keep compressing and move to the right, then we very could well have another healthcare.gov on our hands.

Ms. KELLY. Mr. Thompson, do you disagree with that?

Mr. COOPER. First, overall, we absolutely concur with our GAO colleagues. We are doing some very specific things, including applying learning from healthcare.gov. I would agree with Ms. Cha that a significant failure was the lack of comprehensive end-to-end testing.

We are doing, I think, exactly the right type of approach and it does accommodate the recommendations coming from GAO. First, one way to think about how we are going at CEDCaP is it’s a building block approach. I mean, playfully think Legos. We’re creating modules and we are thoroughly both planning the integration, rigorous architecture around each of those modules. Those modules or building blocks are then being brought together. But we know where the integration of these building blocks occurs, we have both unit testing and system testing for each of these building blocks, and they are being tested again.

So we have full-time. I admit we have to stay on schedule. If we slip schedule, that’s a risk. And I can share what we’re doing to prevent that and to accommodate it if it should occur. But most important, we’ve built into our plan a very aggressive and rigorous iterative approach, agile approach, so that we’re not building out the entire solution before we test anything. We’re testing as we go. The likelihood that we will miss any fatal error becomes significantly reduced by this approach.

Then, when we kind of add and test, add and test, add and test—and when I say add, it’s additional functionality, additional capa-

bility—what we end up with is significantly reduced risk as we move toward our end-to-end testing.

Ms. KELLY. Okay. Mr. Thompson, any other comments?

Mr. THOMPSON. No, I agree with Mr. Cooper. We enjoy the relationship we have with the Department of Commerce and the governance process that we have in place.

I would say one thing, that I was fortunate enough to have been the career person in charge of the 2000 Census. And we successfully delivered a number of very, very complicated IT systems that were state of the art at the time.

When I went to NORC in Chicago we also put in place some very, very sophisticated IT systems. So I have a lot of experience in working with creating IT systems that work. And that's one of the reasons I came back to the Bureau, because I saw an opportunity to realize technology to save the government a significant amount of money and improve accuracy.

Ms. KELLY. According to the operational plans, the Bureau will begin preparations for end-to-end testing of all systems and operations August 17. Is that accurate? 2017.

Mr. THOMPSON. That's accurate. We refer to it as the 2018 test because the Census Day is April 1, 2018, but we have to start in preparations in 2017.

Ms. KELLY. Okay. I yield back also the time I don't have.

Mr. MEADOWS. I thank you.

Mr. Cooper, let me follow up on you indicated that one of your concerns was that it was a proprietary technology, the handheld technology of our previous fiasco, I guess is how I would characterize it, and that somehow because we now have smartphones that that's going to fix the problem.

We had smartphones in 2008, at least we did in North Carolina. I assume you did here in Virginia. So if that's the case, it's not really a hardware problem, it's something a lot more significant than that.

Mr. COOPER. I would agree with you. What I meant by a proprietary solution was is that we were building an integrated hardware platform that was—

Mr. MEADOWS. Who was building it?

Mr. COOPER. Census Bureau and—

Mr. THOMPSON. Harris.

Mr. COOPER. Harris Corporation was the prime contractor on that integrated solution.

Mr. MEADOWS. So when it failed, who pointed the finger at whom?

Mr. COOPER. I would have to defer to my colleagues. I was not present in the Department of Commerce at the time and I don't know the answer.

Mr. MEADOWS. If I see somebody smiling behind you maybe you can ask him.

Oh, you've got that Director Thompson.

Mr. THOMPSON. So I actually was involved in the issue. I was called by Secretary Gutierrez to be on the special panel that give him recommendations, so I did look at it. And it seemed like that there was a lot going on on both sides in terms of defining and agreeing on what the specifications should have been. But the big

problem was they were trying to invent a software—a hardware technology that just didn't exist.

Mr. MEADOWS. But aren't we trying to do—we're trying to develop an app now for an iPhone? I don't want to say iPhone. Android. Whatever it is. I mean, I can see that the endorsement comes. Is that what we're trying to do, is develop an app?

Mr. COOPER. Yeah, the difference—

Mr. MEADOWS. Because here's my concern. We're talking about all this testing that's going on. How do we test something that we haven't even decided what the design is all about. I mean, I don't buy that, that we're testing.

Mr. COOPER. That's fair. In this case what we are already operationally field testing is the following: We are using a well-proven—my remarks—a well-proven platform. Think of this—

Mr. MEADOWS. So we're using the smartphone.

Mr. COOPER. All right. We have developed, it's already developed, we have a developed application—

Mr. MEADOWS. You've got an app.

Mr. COOPER. Yes, sir.

Mr. MEADOWS. So I could download the app on my phone today.

Mr. COOPER. Yes, sir. Yes, sir.

So the field testing, most recently in Scottsdale, Arizona, and it's planned now for several additional broader locations, the application rides on a proven platform rather than that integrated, single hardware integrated with proprietary software that we did in 2010. A very different approach, architected very different.

It allows us to do a couple very important things. First of all, with a proven technology platform, what we are really focused on is, as best we are able to, industry plus best practice, industry best practice, our best practices, secure the platform itself. More importantly, the application is about collecting the data. We now have the ability to separately go after securing the application, access to the application—

Mr. MEADOWS. And I get all that. Let me cut to the chase.

Mr. COOPER. Please.

Mr. MEADOWS. Here is my concern. Looking at your time frame and how it's already moved to the right, and looking at when we will be testing it and when you will go out with your RFP for whether you build it or somebody else builds it, we're running out of time. And if we're sitting there worried about how a handheld device is collecting the very simplest of data, and I come from where Google headquarters, all your data is stored in my district, and so if that's what we are working on at this point, you know, that's like saying we're working on the steering mechanism for a car, believing that the car is going to work okay. It may steer okay, but it may not run.

And so that may be a crude analogy, but we have to be a lot more end to end in terms of what we're trying to do. And my concern is from a technology standpoint, we are nowhere near close, other than conceptual. And I see some heads shaking no, so they're saying that that's inaccurate.

So go ahead, I'll give you this chance to—I'm sorry, I keep reading body language behind you. But go ahead.

Mr. COOPER. We don't have the advantage of seeing behind, but I can see it in the reflection of my glasses.

Mr. MEADOWS. You must not be a parent. We have eyes in the back of our heads.

Mr. THOMPSON. I'm a parent, but my youngest kid is 30 years old.

Mr. MEADOWS. Go ahead.

Mr. THOMPSON. So we have already built a prototype system and tested it in the field and we are——

Mr. MEADOWS. For the data collection?

Mr. THOMPSON. For the data collection and for the control. And we built that system using basically existing technology and existing software from vendors. We only put our parameters into it. We didn't invent new software.

The big decision we have to make is do we scale up this prototype to do the whole Census or do we buy a solution to scale it up. But we already have defined requirements for what we need, that's the beauty of this prototype. So we have the requirements for what the device has to do, and we have engaged the Carnegie Mellon Software Engineering Institute to help us make the right build-or-buy decision and put that——

Mr. MEADOWS. But, Director, how is that different than what we had in 2008? Because we had a defined technology. It was proprietary given that, so maybe it was not proven. But how is it different than that?

Mr. THOMPSON. So in 2008 they didn't have the specifications.

Mr. MEADOWS. So we designed something with not knowing? You've got to be kidding.

Mr. THOMPSON. They were still trying to design the specifications.

Mr. MEADOWS. Okay. All right.

Ms. Cha, let me come to you. What am I missing here in terms of do you have concerns?

Ms. CHA. We do have concerns. And, Chairman Meadows, to be clear, the systems that they're talking about at this time are the applications. These are prototype systems. And, again, in terms of scaling it to the production needs for 2020, we do have concerns that given the available time remaining, that hardening those prototypes so that they meet 2020, that's not an assumption—I think that's a dangerous assumption to make, that these prototype systems will be in place.

But I think for the key lesson learned from 2010, coming out of that, it's that the Bureau underestimated the technical complexity associated with those handhelds. And so even though they are not developing new devices, that lesson is still important in looking at the total magnitude of what they are intending to do with the operational control system to manage the field work with the Internet response option, with the devices that ultimately will be deployed.

When you look at this collectively in the remaining time, I think it's fair for the Bureau to again make decisions now, take steps to reduce scope, take things off of the table so that they are positioned for success.

Mr. MEADOWS. I'm out of time. Let me just share this. We want you to succeed. But the other thing that Mr. Connolly and I and

Mr. Hurd and Ms. Kelly, we don't want egg on our face. And there is no way I'm going to allow this to continue to progress without certainty.

Mr. Cooper, one of the problems I see is the lack of specificity in terms of deadlines and what is to be accomplished by those deadlines. So I'd ask that you work with the committee on providing that.

And ultimately you're going to have four people, as Mr. Connolly said, willing to go and fight for appropriations and get you the tools, but lack of planning and lack of strategic implementation is something that we don't want to find 6 months, 12 months from now where we are spending a lot of money going the other way.

I'm going to recognize the gentleman from Missouri, Mr. Clay, for 5 minutes.

Mr. CLAY. Thank you, Mr. Chairman. And hopefully I can take some liberties with the time just to share with you and the rest of the committee that for the 2010 Census I was sitting in your chair, I chaired the Oversight Committee. And when we ran into the issue with the handheld, it cost the taxpayers quite a bit. It doubled the cost of the 2010 Census compared to the 2000 Census. And it was a mistake that should have been avoided, but wasn't. And so hopefully we have learned from that mistake.

And according to GAO, fundamental weaknesses in key IT management practices contributed to the Bureau not being able to successfully deploy custom-developed handheld enumeration devices for Nonresponse Followup, which increased the cost of the Census by up to \$3 billion.

GAO has reported that the Bureau faces a number of challenges with, "developing and deploying the information technology systems and infrastructure it plans to rely on to conduct the significantly redesigned 2020 Census."

Ms. Cha, is that correct?

Ms. CHA. That's correct.

Mr. CLAY. Mr. Cooper, what is your response to that assessment?

Mr. COOPER. I think, as we have stated, first, we agree with the GAO recommendations. We are working diligently to respond to all of those recommendations. We've shared some of the numbers and the status with you. We'll continue, you have my commitment, we'll continue to keep all of you informed of this status as we move forward.

I heard Chairman Meadows loud and clear over addressing the concern over lack of decision deadlines; same point made by GAO. I hear it. I will certainly personally take the commitment to address it, working with my Census Bureau colleagues.

I think that's the quick summary of the approach, and I'm more than happy to come back, dive into detail, and add any additional information that would be helpful.

Mr. CLAY. All right, okay.

Now, let me address the next one to Mr. Cooper. And let me preface this by saying this is 2015, and if I'm correct, this is the time to ramp up, to get ready for the 2020 Census. And, Mr. Thompson, I heard you loud and clear that this is not the time for Congress to be playing tricks and using smoke and mirrors with the Census Bureau's budget because this is ramp-up time, this is time for you

to be prepared to eventually get to 2020 and conduct the Census successfully and to eliminate the undercounts and the overcounts.

But, Mr. Thompson, in April 2014, GAO reported that the Bureau had not prioritized key IT research and testing needed for its design decisions. How do you respond to that?

Mr. THOMPSON. So what we have done is we have been working very diligently to do research and testing. We started in 2013 when the Bureau—that's before I got there, the Bureau had some budget cuts, and they reprioritized their whole, entire research program. And importantly, they established the key milestone, which was for the beginning of fiscal year 2016, this past September, that they would release an operational plan that did two things. It laid out as many key decisions as could be made and, importantly, laid out a process for making the remaining decisions.

And we have at the Bureau done that. We released an operational plan. It lays out a schedule. It lays out a schedule for making our key decisions, leading up to an integrated end-to-end task in 2018—well, Census Day is 2018, it starts in 2017.

Mr. CLAY. Sure.

Mr. Chairman, may I talk to Mr. Goldenkoff, because he and I were involved in the 2010 Census.

Do you see anything we need to be watching out for going from this time line forward that are similar to what we experienced in the 2010 Census?

Mr. GOLDENKOFF. Yes. I think that both the Census Bureau, the Department of Commerce, and both subcommittees need to be extremely sensitive to the early warning signs. And I think that some of those early warning sides were not heeded back in the lead-up to the 2010 Census. For example, GAO had pointed out a number of issues with the handheld device beginning, I believe, in 2006. There were yellow flags being raised all over the place. And there was just a tendency, I think, for the Census Bureau to perhaps discount some of those concerns.

I think the environment is much better now, the Census Bureau is much more responsive to recommendations from us, to your oversight, from the IG's office, but we just need to be sensitive to those early warning signs. Delays, you know, concerns about the lack of time, they're starting to crop up. So funding issues, need to make sure that the Census Bureau has reliable funding throughout the course of the decade. It is not only the actual amount of the money that they get, obviously that's important, but uncertainty is also an issue too.

So those are some things that right now, while there's still time, we all need to be sensitive to.

Mr. CLAY. Thank you.

Thank you, Mr. Chairman.

Mr. MEADOWS. I thank the gentleman for his insightful questions.

And the chair recognizes Mr. Grothman, the gentleman from Wisconsin.

Mr. GROTHMAN. Thank you. I have a couple of questions here.

Over time I think, you know, insofar as a garden variety legislator who calls about the Census, it's about some of the questions that are asked. Not about the Census every 10 years, but, I don't

know what you call them, these more broad questionnaires that get sent out, correct? What do they call them, the broader—

Mr. THOMPSON. Are you referring to the American Community Survey?

Mr. GROTHMAN. Yeah. How long do you retain the information on those surveys?

Mr. THOMPSON. Since they are part of the decennial Census, we retain a permanent record of them. And 72 years after they're collected then they can be made available to the American public.

Mr. GROTHMAN. You mean, you're even going to make those available sometime, even 72 years?

Mr. THOMPSON. In 72 years, yes, sir.

Mr. GROTHMAN. What are we doing to make sure? I mean, some of those things when I get questions they wonder why is it any of the government's business. And while it hasn't necessarily been your organization, certainly other government organizations have not been able to protect their data. Maybe that's an impossible thing to do, but I am just saying you've got to realize that when you fill out a government form there's a possibility that someday the whole country will know what's on it.

Are you doing anything—I'll put it this way. On those community surveys, is there any reason why that stuff has to be held more than 3 or 4 years?

Mr. THOMPSON. Yes, sir, it's part of our national records program. And so as being part of the decennial Census we're mandated to make those records available to the public for a number of important reasons, so people can look at—can do ancestor research, so people can understand how their ancestors were working.

Mr. GROTHMAN. I understand it's kind of cool that I can look up and find my, whatever, great-grandfather was living wherever in 1880. But beyond that, there's stuff here that people don't want to have let out. So I'll ask you again, why—given that I would assume sooner or later you're going have a breach of your database, every other government agency seems to—is there any reason why you personally feel that if I fill out a form some information of which may be considered somewhat personal or at least they don't want everybody to know, any philosophical reason why that stuff has to be kept around?

Mr. THOMPSON. As I said, we at the Census Bureau believe it's very important that we maintain a record of the Census and make that available in 72 years.

Now, the Census Bureau takes security and privacy of our information very, very seriously, and we could go into great lengths about how we go about protecting the PII of our respondents and how we have a number of layers between that and any way to get to it. And so that's job one. Plus, we have some really severe penalties in place if there is willful disclosure of any kind of information.

Mr. GROTHMAN. Well, I'm sure that the IRS takes their responsibility seriously and whatever our equivalent of human resources is takes their responsibility seriously. And I realize we want to keep the bare bones minimum of the Census available out there. Like I said, we should maybe always know that Glenn Grothman in

2020 is living in Glenbeulah, Wisconsin. But why are we keeping all this more personal stuff? What is the sense of keeping that available for either some hacker to get ahold of it or even some busybody to look at it in 70 years?

Mr. THOMPSON. Congressman, we are actually mandated by law—and don't ask me the specific cite, I'll be happy to get it to you—to maintain those records as part of the decennial Census.

Mr. GROTHMAN. Okay. But you don't have an opinion on it personally?

Mr. THOMPSON. My opinion is to do the Census as the Congress shall direct. So that's basically my opinion.

Mr. GROTHMAN. Okay. I'll give you one more question here, which is kind of the parting question. We still don't know whether you're going to use the NDNH or the kids link sources. When are you going to decide whether you're going to use those or not, or what is going to determine whether or not you're going to use them, have access to them?

Mr. THOMPSON. So we're going to make our final decision on the exact administrative records we'll use, which will be in 2018, and that would include whatever we've been able to put together, and that's what we're going to go forward with. We're not going to change after that point.

Mr. GROTHMAN. Thank you.

Mr. MEADOWS. I thank the gentleman.

The chair recognizes the gentlewoman from New York, Mrs. Maloney, for 5 minutes.

Mrs. MALONEY. I thank the gentlemen and ranking member for calling this important hearing and all of the participants today. The Census is mandated in the United States Constitution. It's a requirement illustrated right in the Constitution. And historians constantly refer to it, documents refer to in it ways that help us understand our history and understand our country. And I'm glad that we're focusing on it, because if we don't have good data, if we don't have a good Census, then we don't have good policy.

I would also say the Census is relied on more and more by businesses and others to understand our country, where we are, where we're going, what our needs are. And I would say the allocation of Federal funds is dependent on Census data, as is oftentimes public policy in general, and it's incredibly important.

But that said, I want to really associate myself with the comments of the prior speaker on the importance of keeping it confidential, that this is a critical trust that this be kept confidential until 70 years later. And I do want to cite that there happens to be a Broadway play playing right now in New York on a violation of the Census data during World War II, where the Census data was used to round up Japanese Americans and place them in detention camps. It's probably the worst chapter of the Census in the history of our country, that our sacred data that we pledge will be confidential was violated. So that was a bad, bad, bad era.

But I do want to go back to cybersecurity and that we have to take the cybersecurity seriously. And I want to say that in light of a recent cyber incident at the Bureau, I want to better understand the data controls currently in place. I don't believe that the Census would ever give up data in this day and time, you know, willingly.

I don't believe that at all. I think it's totally confidential. But I am concerned about a cyber incident.

And I'd like to ask Mr. Thompson, the Director, on July 24 you issued a statement confirming that the Bureau, and I am quoting from you, "experienced an attack to gain access to the Federal Audit Clearinghouse." And can you first explain to us what is the Federal Audit Clearinghouse? And your statement went on to say, "While our IT forensics investigation continues, I want to assure you that at this time every indication is that the breach was limited to the this database and that it did not include personally identifiable information provided by people responding to our Census and surveys."

So your comment please on that.

Mr. THOMPSON. Thank you. So that is a database—what we call a Web-facing database that we maintain that sits between our firewall and the outside. And so that database was maintained for people to put certain types of information in which is nonsensitive information and that's the database that, unfortunately, got breached. We've taken steps since then, which Mr. Cooper could explain better than I can because he's more of an IT person, to rectify that problem.

But importantly, we also have safeguards that will not let anything past our firewall which will get to the personally identified information at the Census Bureau, and that was not breached in that circumstance.

Mrs. MALONEY. Okay. Can you share with us the results from the forensic investigation, Mr. Cooper?

Mr. COOPER. Yes, ma'am, I can.

Mrs. MALONEY. And what steps have you taken to prevent a similar cyber incident in the future?

Mr. COOPER. Okay. First what our forensic analysis has shown to date is that gives us the time line of the attack and the source addresses from which the attack was launched, both—

Mrs. MALONEY. Where do you think it came from? Where do you think the attack came from?

Mr. COOPER. In this environment, could I maybe follow up and get with you all in a different physical environment to discuss that with you please?

Mrs. MALONEY. Okay. Thank you.

Mr. COOPER. If you'd allow me to do that.

Now, we've also worked very, very closely with our colleagues at the FBI and with the Department of Homeland Security, so that that becomes part of the intelligence community environment, that helps us at the Census Bureau, it helps our colleagues across the Federal Government. So there's the first part that I can answer in this setting today.

Second part. We have indeed installed new security measures, and let me give you a quick summary of those.

First of all, we have—we're in the process of implementing Web application scanning capability that we did not have broadly in place at the time of that breach. We've also included funding in our fiscal year 2016 budget to bring the cybersecurity company we used in fiscal year 2015, "we" in this case being the Census Bureau, back to do another scan.

And I have instituted as the CIO regular scanning. We have reached out to some other Federal departments to assist us with what's called blue team assessments, blue team attacks, which are friendly attacks against our environment to help identify threats and vulnerabilities that we may not have uncovered ourselves.

We're also actively engaged with our colleagues at the Department of Homeland Security. We have fully implemented all available Einstein precautions to enhance and strengthen our cyber scanning capability to identify threats and critical vulnerabilities. We're working very closely with the Department, Department of Homeland Security, to implement as part of what's known as phase two of the Continuing Diagnostics and Mitigation program of DHS. We are now working and we will be deploying additional CDM capability, more scanning and software tools, sensors, in our networks.

All of this strengthens our environment and enhances our capability from the time of the breach even till now, and will continue to do that going forward.

Mrs. MALONEY. Well, my time has expired, but possibly you could get back to us in writing the controls the Census Bureau has in place to protect this collected data.

Mrs. MALONEY. I think it's very important. And people want to know that their data is protected, as the gentleman mentioned earlier.

I thank you and I yield back.

Mr. MEADOWS. I thank the gentlewoman.

Just for planning purposes, we're going to have a very brief second round of questions, and by very brief only a few members, I think, are going to participate in that.

So the chair will recognize the chairman of the subcommittee, Mr. Hurd, for a series of questions.

Mr. HURD. Gene Dadaro is going to get mad at me for asking this question, but, Ms. Cha, I hear there is a position available at the Bureau, the Census Bureau, for chief of business integration. Would you be interested in taking a sabbatical to take that position? I think you would be well-positioned for that.

But, Mr. Cooper, Mr. Thompson, if I understand this correctly, you're basically crowdsourcing people to answer some—you know, to get better information, have them get that information into you. We didn't even get to ask many questions. Not knowing how many people you expect to respond to that, that's going to drive your technology decisions and the infrastructure. Two years out, that's difficult.

But let me understand this bring your own device or buying a new device. Are we talking about having—buying actual smartphones that have the app already on there with the security systems established or are we talking about a new designed device that's not a smartphone but it's used specifically for you all? I'm confused.

Mr. COOPER. Okay. In this case it would be the former of your two scenarios.

Mr. HURD. Okay. So it's a smartphone that everyone is going to use. And when are you going to make the decision—

Mr. COOPER. That would be option one, and that would be a government-furnished device. Options two would be bring your own device and we will then load the—

Mr. HURD. And when are you making this decision on which?

Mr. COOPER. September 2016.

Mr. HURD. September—how many months is that?

Mr. COOPER. It's about not quite a year from now.

Mr. HURD. Why 12 months, why that long to make a simple decision that's going to base—that's going to drive your entire plan?

Mr. COOPER. That allows us to complete the planned and already in motion, in process set of field operation tests so that we can make both an economic-based determination as well as a security-based determination and include privacy and functionality.

Mr. HURD. I recognize the privacy and security concerns of having people bringing their own device and uploading that information onto some system or server that touches all of your databases. But this is not a new technology, right? Everybody on this platform have probably block walked at some point in time and this is just block walking on steroids.

And I represent a very rural part of Texas that doesn't have cell phone service and we would still be able to do that. So the necessity to have a year of testing a technology and a process that is used pretty significantly is shocking to me. And that's something that I'd love to continue to talk about another time.

I have two final questions, a separate issue. It is my understanding that missionaries and other individuals temporarily overseas with a clear intent to return would not be included in the Census. Is this correct?

Mr. THOMPSON. Congressman, we are in the process right now of determining the residence rules that we will use for the 2020 Census. Recently we sent out the rules that we use for where to count people, who to count in 2010 for public comment. We received a number of comments, and we are now in the process of responding to those comments, and we will then issue our plan for 2020. In the last Census, if someone was permanently away from the United States we did not count them in the United States.

Mr. HURD. Okay. We would welcome your documentation in writing on this process, it is important to a number of members.

Mr. HURD. And also, in the 112th Congress, Senator Hatch and Representative Bishop introduced companion bills requiring that all citizens of the United States temporarily living abroad at the time of the Census be counted and attributed to the State they had most recently lived in. This legislation also provided for the use of administrative records to assist in the count. Can you describe any concerns you have with this legislation? And we would love to see that in writing if you all haven't done the analysis of that.

Mr. THOMPSON. At this point I'd have to respond to you in writing.

Mr. HURD. Thank you very much.

I yield back, Chairman. Thank you.

Mr. MEADOWS. I thank the gentleman.

The chair recognizes the ranking member, Mr. Connolly, for a series of questions.

Mr. CONNOLLY. Thank you.

Mr. Cooper, these are, I hope, brief answers to brief questions. Following up on Mrs. Maloney's questioning, you mentioned Einstein. Part of the problem with the breach at OPM was Einstein 3, I believe, was not yet fully implemented. Are we at Einstein 3 for the Census?

Mr. COOPER. We are where our telecom providers have already put the Einstein 3 capabilities in place.

Mr. CONNOLLY. Good.

Mr. COOPER. There is one telecom provider that is not quite at Einstein 3.

Mr. CONNOLLY. Okay. Good to know.

Mr. COOPER. The answer is yes.

Mr. CONNOLLY. Secondly, in response to, I think, Mr. Blum's questioning, I think you acknowledged that there could be roughly 100 different systems—I guess at Census or Commerce?

Mr. COOPER. That would be Census.

Mr. CONNOLLY. Census. Presumably some of these are legacy systems.

Mr. COOPER. Absolutely, yes. I don't know the quantitative breakout. I can follow up and—

Mr. CONNOLLY. Some of those systems do not lend themselves as a result to encryption. Is that correct?

Mr. COOPER. That is correct, in some cases.

Mr. CONNOLLY. So that makes them more vulnerable, not less vulnerable to hacking.

Mr. COOPER. It's my understanding—and I'm going to defer, I'm going to ask over my shoulder to kind of make sure I don't misstate this—it is my understanding that while we do have legacy systems involved somewhere in that 112, it's my understanding that we don't have a situation where a legacy system that's directly involved in supporting Census operations leading to the 2020 decennial is not able to be encrypted.

Is that a valid statement?

Okay. Yes. So that would not be a situation we would run into in support of the 2020 Census.

Mr. CONNOLLY. Okay. Thank you.

I think that's it for now, thank you very much.

Mr. MEADOWS. Well, I want to close by thanking each of the witnesses here. I guess the most important thing, I was talking to the ranking member, what we would like to do is not in a hearing setting, but necessarily in a briefing setting is set some regular updates of which actually all four of you would be welcome for those regular updates. I think over the next 12 months I see that as a critical window. And as much as we want to talk about 2020, I think the next 12 to 18 months is our go-or-no-go timeframe. And so in doing that, to have quarterly briefings as to where we're making progress and where we're not.

Is everybody okay with doing that?

I see—let the record show everybody nodded in the affirmative.

And, Director Thompson, Mr. Cooper, thank you, thank your staff for the work. I mean, sometimes it is only thought about when we get the results or when we have someone knocking at our door. There is obviously years of planning that go ahead of that time to

make sure that it's done seamlessly. And we're going to count on 2020 being done seamlessly. And so I would do that.

I would also be remiss in saying that many of the recommendations that the GAO are making are raising concerns. If they're not happy, I'm probably not going to be happy. And so I just want to stress that, that it is critical that we work hand in glove together, because failure is not an option.

And with that, thank each you for your testimony here today.

If there is no further business, without objection, the subcommittee stands adjourned.

[Whereupon, at 4:16 p.m., the subcommittees were adjourned.]

APPENDIX

MATERIAL SUBMITTED FOR THE HEARING RECORD

Question for the Record
 Submitted by: Congresswoman Plaskett
 Hearing on "Preparing for the 2020 Census: Will the Technology be Ready?"

Importance of Securing Accurate Census Data From All US Territories Through The American Community Survey

Thank you, to our witnesses for being here today. You play an important role in ensuring the successful accounting for all people within this country.

While every ten years, key data is collected from our nation's population through the Census, the American Community Survey provides vital information on a yearly basis about our nation and its people. Information from the survey generates data that help determine how more than \$400 billion in federal and state funds are distributed each year. That information is used to make critical decisions affecting the operation of our government and private industry.

My district along with Guam, American Samoa, and the Northern Mariana Islands are not included in the American Community Survey and without it, there is a misrepresentation about jobs, occupations, educational attainment, veterans, home ownership, and economic growth in our districts.

With the closure of HOVENSA refinery, thousands of citizens have left the US Virgin Islands, economic development projects stalled and broader borders stunned; my district has suffered a major economy downturn.

This single company closure has led to millions of federal dollars leaving the Virgin Islands and an increase in local poverty numbers.

With such key decisions riding on the accuracy of Census data, it is critical the Bureau work with my office to find ways to increase funding so that all territories are included in the American Community Survey or develop a separate survey specific to the US Territories.

With all the fore mentioned stated, Mr. Thompson,

Q: As the Director of the Census Bureau, what can you do to assist the US Virgin Islands in being included in the American Community Survey?

Q: Do you agree that including some territories (Puerto Rico and the District of Columbia) and not all is a disservice to some US populations?

Your testimony states that funding is of major concern at the Bureau

Q: Can you explain how I might be able to work with my colleagues in Congress to assure that your Bureau's funding is not cut?

Questions for The Honorable John H. Thompson
 Director
 U.S. Census Bureau

Questions from Chairman Mark Meadows
 Subcommittee on Government Operations
 Hearing: “Preparing for the 2020 Census: Will the Technology be Ready?”

1. In October 2015, the Bureau released its Operational Plan for the 2020 Census.
 - Can you describe specific steps the Bureau plans to take from here on out to improve the Bureau's scheduling practices and to meet the 2018 “dress rehearsal” deadline?

The Census Bureau is fully committed to ensuring that the 2020 Census planning and development process is on time and on budget. To that end, we have taken or are taking a number of actions, including delivery of our operational plan three years earlier than we did for the 2010 Census. The 2020 Census Operational Plan outlines 350 major design decisions for the next census, 168 of which have already been made. It puts into place a process to make remaining decisions that will allow for adequate testing and time to prepare for the 2018 Census End-to-End Test. More details about each of these decisions can be found in the Operational Plan at this link: <http://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/planning-docs/operational-plan.html>. The dates for making these decisions are included in both the operational plan and the 2020 Census Lifecycle Schedule. This integrated master schedule facilitates the integration and coordination across census tests and operations. Additionally, by the end of March 2016, the Census Bureau will baseline the integrated schedule for the 2018 Census End-to-End Test.

The Census Bureau has a number of program management efforts underway for the 2020 Census program that also will help ensure readiness for the 2018 Census End-to-End Test. For example, , the 2020 Census Program has categorized the work of the 2020 Census in alignment with the enterprise work breakdown structure that is currently being developed by the Census Bureau. The 2020 Census Program work breakdown structure is as follows:

1. Program Management
2. Census and Survey Engineering
3. Frame
4. Response Data
5. Publish Data
6. Test and Evaluation
7. Infrastructure

*The 2020 Census Program will ensure that scope definition, cost estimation, budget formulation and execution, and **schedule development and management**, are conducted using these categories to organize work. This follows best practices outlined by the*

"Government Accountability Office Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Capital Program Costs" and the "Government Accountability Office Schedule Assessment Guide."

The 2020 Census Program has successfully used this model to conduct the 2020 Census Research and Testing Program over the past four years, including the following tests:

- 2012 National Census Test;
- 2013 National Census Contact Test;
- 2013 Census Test;
- 2014 Census Test;
- LUCA Focus Groups;
- 2014 Human-in-the-Loop Simulation;
- 2015 Address Validation Test;
- 2015 Optimizing Self-Response Test;
- 2015 Census Test; and
- 2015 National Content Test.

To support the 2020 Census schedule development and maintenance, the Census Bureau has an office devoted to this work. This staff includes a chief and four federal employees. Through a schedule support contract, we also have 15 contractors certified in schedule development, two consultants, and two certified business analysts. All of these schedulers are embedded within the 2020 Census Program project teams to create the schedule, status the schedule, evaluate the impact of change requests on the schedule, and monitor progress on a regular basis. The Decennial Census Management Division managers meet weekly to ensure the schedule is on track. Additionally, this entire group is working closely with the 2020 Census Program budgeting staff and the Census Bureau's Office of Cost Estimation, Analysis, and Assessment, to ensure a fully integrated schedule and budget system.

We also have a formal and rigorous risk management process in place at the program and project level, which takes into account both schedule and budget. This includes developing mitigation plans for all risks, and contingency plans where appropriate. This includes mitigation plans for several risks that relate to the possibility of delays, late design changes, and major budget shortfalls.

Finally, to ensure transparency within the 2020 Census Program, there are several mechanisms used to keep stakeholders informed, such as a Quarterly Program Management Review to brief both internal and external stakeholders, which we began webcasting at the beginning of this year. Additionally, we brief the Department of Commerce, Office of Management and Budget, the Government Accountability Office, the Office of the Inspector General, and Congress, on a regular basis.

All of these are important steps to ensure we stay on schedule in preparing for the 2018 Census End-to-End Test, and for the 2020 Census itself.

- What is the number of major decisions you need to make and how many of these have you made? If there are decisions yet to be made, why is this the case?

The 2020 Census Operational Plan outlines 350 major design decisions for the next census, 168 of which already have been made. These decisions were data driven, based on information collected during the 2020 Census Research and Testing Program, spanning from Fiscal Year 2012 through 2015 and from the American Community Survey. During this period, we conducted several tests, including the:

- 2012 National Census Test;
- 2013 National Census Contact Test;
- 2013 Census Test;
- 2014 Census Test;
- LUCA Focus Groups;
- 2014 Human-in-the-Loop Simulation;
- 2015 Address Validation Test;
- 2015 Optimizing Self-Response Test;
- 2015 Census Test; and
- 2015 National Content Test.

The focus of these tests was concentrated on the four key innovation areas associated with the reengineered concept of operations for the 2020 Census – Reengineering Address Canvassing, Optimizing Self-Response, Utilizing Administrative Records, and Reengineering Field Operations.

There are 182 decisions that still need to be decided, 167 prior to the 2018 Census End-to-End Test. We have carefully planned the dates for making each of the decisions, with the goal of conducting a successful end-to-end test. The dates for these forthcoming decisions are detailed in the operational plan which lays out the path to the end-to-end test and to the 2020 Census. More details about each of these decisions can be found in the Operational Plan at this link: <http://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/planning-docs/operational-plan.html>. These decisions will also be data driven, based on information collected from upcoming tests and operations, including the:

- 2016 Census Test;
- 2016 – 2019 annual Master Address File (MAF) Coverage Study;
- Address Canvassing Test;
- 2017 Census Test; and
- Several specialized tests focused on Puerto Rico, group quarters, and coverage measurement.

The focus of these tests will move from researching and testing prototypes and proofs of concept, to operationalizing and implementing an integrated suite of operations and systems. These tests are critical to finalizing the detailed operational plans for the 2020

Census prior to the 2018 Census End-to-End Test.

Examples of questions to be answered include:

1. *How will the field reengineered concepts tested for NRFU be used for In-Field Address Canvassing?* This will be tested and answered as part of the 2016 Census Test and will be answered by January 2017.

2. *What are the optimal designs for the questionnaires (including size and page layout) and non-questionnaire materials for the 2020 Census?* This will be tested as part of the 2016 and 2017 Census Tests and initial decisions will be made October 2017 and final decision in August 2018.

- Please add perspective by comparing your current state of preparation for the upcoming census to past censuses, and please provide supporting documentation that is relevant.

In 2013, in response to funding constraints, we reprioritized the 2020 Census Research and Testing Program. At that point, we established the end of Fiscal Year 2015 as a key milestone for releasing the 2020 Census Operational Plan. On November 3, 2015, Director Thompson informed the Subcommittees that we had met that goal, and had publically released the plan on October 6, 2015; three years earlier than for the 2010 Census cycle. In comparison to last decade, we released this plan three years earlier than the corresponding plan for the 2010 Census. Releasing our plans now enables us to communicate our plans and decision-making process to Congress and other stakeholders three years earlier as well. In addition, the operational plan defines 34 operations critical to success. For each of those operations, the operational plan documents the purpose, a description of the operation, lessons learned from the 2010 Census, opportunities to innovate, work completed (decisions made), decisions yet to be made, cost and quality impacts, and risks. As related to schedule, the operational plan highlights key milestones for each operation, in alignment with the 2020 Census Lifecycle Schedule. As related to budget, the operational plan documents a lifecycle cost estimate of \$12.5 billion, avoiding over \$5.2 billion in costs as compared to conducting the 2020 Census using the same methodology as was used in the 2010 Census.

On a quarterly basis, the Census Bureau conducts a publicly available Program Management Review where the focus is on the scope, schedule, resources and status of the 2020 Census program. All of this information, including the 2020 Census Operational Plan and other supporting documentation, is available on our website at this link: <http://www.census.gov/2020census>

2. In Ms. Cha's statement, she states that by August 2017, the Bureau intends to begin preparations for end-to-end testing to validate that CEDCAP systems are ready to go live on Census Day. She states there are critical IT decisions yet to be made, including whether the Bureau will build or buy the needed systems. What is the Bureau doing to offset the risk Ms.

Cha laid out of potentially not having all systems in place in time for end-to-end system and operations testing?

The Census Bureau has an extensive risk mitigation strategy. We have a schedule of when critical IT decisions need to be made for the 2020 Census. The schedule for the 2020 Census has been integrated with the CEDCAP schedule to assure that the critical functions needed to support the Census are delivered on time. A key part of our risk mitigation strategy has been to build working prototypes and test them throughout the decade. This process allows us to ensure that capabilities and requirements are identified and documented early. We have successfully fielded several modules of CEDCAP, and CEDCAP modules are in production in other Bureau programs. In addition, as stated in the 2020 Census Operational Plan, we are making innovative use of existing technology and software instead of inventing on our own. We developed a working prototype for the 2020 Census that we successfully tested in 2015. Based on this prototype, we have incorporated most of the 2020 Census Program specifications into the next iteration. We will finish this work based on the results of tests in 2016 and 2017 in time for the 2018 End-to-End Test.

Recognizing the need for timely decisions on which Census Enterprise Data Collection and Processing (CEDCaP) systems we need to build internally, and which we need to buy from external sources, we have engaged the Carnegie Mellon Software Engineering Institute to guide us and to ensure that the decision making process is successful. We created a technical design decision framework that determines key design capabilities based on business capabilities already identified in the 2020 Census Operational Plan, and helps conduct Capability Analysis and Assessment of in-house and Commercial-Off-The-Shelf (COTS) solutions.

Based on our review of the GAO concerns regarding the timing of our IT decisions, we conducted a baseline Fiscal Year 2016 assessment. The Fiscal Year 2016 System Readiness Baseline Assessment was completed on January 12, 2016. It was then presented to, and approved by Census Bureau executive leadership on January 19, 2016. A summary of the review also was presented at the January 2016 Program Management Review, which was attended by GAO and other oversight groups and stakeholders. The changes to the 2016 Census Test business and capability requirements resulting from the assessment were identified by the Census Bureau, and IT system teams are working to complete the implementation of the baseline of business and capability requirements. For the Address Canvassing Test, the 2017 Census Test, and the 2018 End to End Test, the Census Bureau also reviewed and updated the business and capability requirements. The Census Bureau then assessed existing systems for their ability to fulfill the business and capability requirements.

The combination of the work that Carnegie Mellon is conducting, with the Capability Analysis and Assessment, will determine which systems will be built or bought, by March 2016. We will build out the chosen systems (COTS solutions need customization) and test them (whether they are built in-house, or COTS, or a combination) in the 2017 Census Test and 2018 Census End-to-End Test. Following the 2017 Census Test, we will revise systems

as needed and use them in the 2018 Census End-to-End Test, implementing the planned census operations in real-time. This approach helps mitigate the risks detailed in Ms. Cha's statement.

3. Given the time remaining, has the Bureau considered removing any aspects of the redesign or future testing in order to improve the chances of successfully implementing the most critical aspects of the redesign, such as the operational control system? If so, when does the Bureau believe is the appropriate time to make those decisions?

The Census Bureau has been working closely with GAO to understand and respond to its recommendations. In order to speed up decision making on certain remaining decisions, the Census Bureau conducted a careful assessment of scope related to the 2016 Census Test, Address Canvassing Test, 2017 Census Test, and most importantly, the 2018 Census End-to-End Test, following the release of the 2020 Census Operational Plan in October 2015. We have reviewed the objectives and refined the scope of each test to ensure that sufficient resources will be devoted to planning and implementing the operations and systems needed to successfully conduct the 2018 Census End-to-End Test. Towards reducing risks to operations and schedule, the Census Bureau then made several decisions regarding the scope of the 2020 Census program. For example:

1. *We will move forward with the use of Device as a Service for the 2016 Census Test and the 2020 Census. We have awarded a service contract for the 2016 Census Test that will acquire mobile devices (smart phones and tablets), determine the best solution for wireless services, and handle device management logistics. We will release an RFI for Device as a Service in support of the 2020 Census in the Spring of this year. We will not move forward with the use of Bring Your Own Device for the 2016 Census Test or the 2020 Census. While our testing proved that BYOD is technically feasible, we identified several significant risks that could impact the census. A good example is having to manage updates to multiple mobile device operating systems days before enumeration begins.*
2. *We will move forward with the use of cascaded training for our field staff, meaning our supervisors will train their employees. We found that the use of master trainers, people especially trained in delivering training, did not prove overly effective during the 2015 Census Test. Additionally, we believe that our field staff will benefit from the face to face contact and interaction with their supervisors prior to conducting field work.*

Using the refined scopes as a guide, we conducted the Fiscal Year 2016 System Readiness Baseline Assessment that was completed on January 12, 2016. It was then presented to, and approved by, the 2020 Census Program executive leadership on January 19, 2016. A summary of the review also was presented at the January 2016 Program Management Review, which was attended by GAO and other oversight groups and stakeholders. The Census Bureau has identified the changes to the 2016 Census Test business and capability requirements resulting from the assessment, and IT system teams are working to complete the implementation of the baseline of business and capability requirements. For the Address Canvassing Test, the 2017 Census Test, and the 2018 End to End Test, the Census Bureau also reviewed and updated the business and capability requirements. The Census

Bureau then assessed existing systems for their ability to fulfill the business and capability requirements.

The combination of the work that Carnegie Mellon is conducting, with the Capability Analysis and Assessment, will determine which systems will be built or bought, by March 2016. We will build out the chosen systems (COTS solutions need customization) and test them (whether they are built in-house, or COTS, or a combination) in the 2017 Census Test and 2018 Census End-to-End Test. Following the 2017 Census Test, we will revise systems as needed and use them in the 2018 Census End-to-End Test, implementing the planned census operations in real-time.

It is important to mention that all decisions are contingent upon the receipt of adequate funding. Otherwise, the Census Bureau will need to reprioritize our activities to ensure the systems are built. With the receipt of the Congressional budget each year, we will continually reassess and reprioritize.

4. For the following key IT decisions, what is the Bureau doing to reassess and reprioritize the timing of these decisions to ensure that systems are in place for end-to-end testing, or to mitigate the risks associated with making these decisions later?
 - a. The extent to which BYOD and Device-as-a-Service will be used to support field operations – January 2016

The Census Bureau's executive leadership recently made a decision that a Bring Your Own Device (BYOD) strategy for provisioning equipment to enumerators will not support the 2020 Census. Instead, the Census Bureau will implement the Device as a Service strategy for provisioning equipment to enumerators in upcoming census test and in the 2020 Census.

Following the release of the 2020 Census Operational Plan in October 2015, the Census Bureau undertook a rigorous planning session to outline the execution details of the remaining tests and examine where operational risk could be reduced. As part of this effort, the Census Bureau considered GAO's recommendations to examine whether any decisions could be made ahead of schedule to reduce risk. This most recent decision is the result both of responding to GAO's recommendation to focus the program, but also importantly because remaining open questions must be resolved prior to the end-to-end test scheduled to begin in 2017. We must prioritize areas in the 2016 and 2017 Census tests where test data is insufficient to make final decisions.

The decision to narrow the focus to Device-as-a-Service is a risk-based decision that included a review of the experiences from prior tests in 2014 and 2015. While the use of a BYOD option did not generate any observable concerns from respondents during these tests, the program faced a number of challenges that present operational risk to the 2020 Census:

- *A BYOD option would have resulted in a diverse universe of smartphones that would have included numerous providers and operating systems over which the Census Bureau had no control. Managing updates to mobile device operating systems days before or during field operations (when hours of delay can cost millions of dollars of wasted productivity and reduce stakeholder confidence in mobile device data collection);*
- *Development of an acceptable use policy that temporary employees would agree to would be difficult and time consuming; and*
- *Development of a data charges reimbursement policy that would compensate temporary employees for additional costs related to their employment while preventing fraud. Separating Census Bureau related charges from personal charges would have caused unreasonable time delays and increased paperwork processing for employees and the Census Bureau.*

We made this decision in January 2016. Documentation can be found here:

http://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/memo-series/2020-memo-2016_01.html

b. Plans for mobile devices use— January 2016

As documented in the 2020 Census Operational Plan, the Census Bureau has already decided to use mobile devices during the 2020 Census in support of address listing, enumeration, work availability reporting, and time and expense reporting. As mentioned just above, in January 2016 we made the decision to implement the Device-as-a-Service strategy for provisioning equipment to enumerators in upcoming census tests and in the 2020 Census. We are currently documenting final decisions for planned uses of mobile devices.

Completed in January 2016.

c. Build or buy decisions for each 2020 capability – June 2016

Recognizing the need for timely decisions on which Census Enterprise Data Collection and Processing (CEDCaP) systems we need to build internally, and which we need to buy from external sources, we have engaged the Carnegie Mellon Software Engineering Institute to guide us and to ensure that the decision making process is successful. We created a technical design decision framework that determines key design capabilities based on business capabilities already identified in the 2020 Census Operational Plan, and helps conduct Capability Analysis and Assessment of in-house and Commercial-Off-The-Shelf (COTS) solutions.

We currently plan to make the decisions on build versus buy no later than March 31, 2016.

d. Projected demand on IT infrastructure—June 2016

Working with systems engineering experts from the MITRE Corporation, we are developing seven external demand models (highlighted in bold) that project volume and patterns of critical 2020 Census operations, including:

- The number of **applicants** for management and field staff positions;
- The **onboarding** of managers and field staff;
- The number of field staff requiring “help desk” support from the **Decennial Service Center**;
- The number of tribal, state, and local government partners requiring “help desk” support from the **Geographic Partnership Support Desk**;
- The number of respondents answering the 2020 Census questionnaire using the **Internet Self-Response Option**;
- The number of respondents answering the 2020 Census questionnaire by phone using the **Census Questionnaire Assistance Self-Response Option**; and
- The number of respondents answering the 2020 Census questionnaire using the **Paper Self-Response Option**, resulting in the need to conduct data capture operations.

Additionally, as we learn from the census tests, we are designing our operations to help “smooth” or “even out” volume and patterns. For example, a key design decision documented in the 2020 Census Operational Plan is that when we invite the population to self-respond using the Internet, we will use a staggered mail approach. The population will receive mailed invitation materials on a staggered basis instead of all at once. This will help us manage the call volume coming into Census Questionnaire Assistance.

We currently plan to make the decisions on projected demand on IT Infrastructure by June 2016.

e. Solutions architecture for the 2020 census – September 2016

The preliminary 2020 Census solutions architecture was delivered on September 30, 2015. This included the project-level business requirements and capability requirements, including the associated business process models, for the thirty 2020 Census operations (out of 34) that require solutions.

Lessons learned during the 2016 Census Test will allow us to refine subsequent versions of the solutions architecture for the 2020 Census by September 2016. As with the 2020 Census Operational Plan, the solutions architecture will be continually refined throughout the census cycle.

We currently plan to make the decisions on solutions architecture by September 2016.

f. Methodology for conducting non-ID response validation – September 2016

In the past year, we have tested two distinct methods for validating responses received from households or individuals that did not provide a Census ID. We recently received recommendations from the JASON Advisory Group about additional methods to explore, and we plan to test those over the next two years in order to determine what we will implement for the 2018 Census End-to-End Test and for the 2020 Census.

Our 2020 Census Operational Plan provides additional details on these decisions planned for September 2016.

g. Response rates for all self-response modes – October 2017

As documented in the 2020 Census Operational Plan, we currently project a 63.5 percent overall self-response rate after six weeks of self-response, with 47 percent via Internet, 5.3 percent via telephone, and 11.2 percent via paper. We are designing the Internet data collection system to handle this response load. We will use results from the tests in 2016 and 2017, plus information from external demand models, to refine these estimates if necessary.

We currently plan to refine these estimates by October 2017.

h. How the Bureau's address and geographic databases will be used to support reengineered field operations – October 2017

As outlined in the 2020 Census Operational Plan, the Census Bureau's Master Address File/Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) System will be used as the foundation to support the concept of reengineered field operations. These databases will serve as the base for our geospatial review during In-Office and In-Field Address Canvassing, key components of our reengineering efforts. Staff will validate and change addresses within the MAF as necessary, as well as add newly identified addresses. The TIGER spatial database will be used to depict the network of streets, highways, railroads, streams, and political, administrative, and statistical boundaries.

We will use information and lessons learned from the 2016 Census Test, the Address Canvassing Test, and the 2017 Census Test, to finalize how the MAF/TIGER System will be used to support the reengineered field operations for the 2020 Census.

We currently plan to make these decisions by October 2017.

i. Expected scale of 2020 census non-ID workload – September 2018

While the 2013 – 2015 Census Tests have demonstrated that most people will respond using the unique Census identification code (Census ID) provided, the Non-ID Processing operation is focused on making it easy for people to respond to increase self-response

rates. The operation accomplishes this by:

- Providing response options that do not require a unique Census ID;
- Maximizing real-time matching of Non-ID respondent addresses to the Census Bureau's Master Address File;
- Accurately assigning nonmatching addresses to census blocks; and
- Conducting validation of all Non-ID responses.

Results of the 2013 – 2015 Census Tests have informed the Census Bureau's projection that 10 percent of the housing units, approximately 14 million, will respond via Non-ID. Future tests will give us a further set of data points for workload modeling efforts. By pursuing a cloud solution for real-time Non-ID processing (where the large majority of Non-ID cases are processed), we can dynamically scale the processing to meet increased demand, thus addressing any potential risk of a larger-than-expected workload. Results from testing the cloud solution in 2016 and 2017 will provide us with sufficient information to inform IT infrastructure planning for the 2018 Census End-to-End Test and for the 2020 Census.

We currently plan to update our estimate of this workload by September 2018.

5. Please describe your contingency plans if the IT systems have not been fully implemented or integrated in time for the 2020 Census.

For the 2020 Census, the Census Bureau has a risk management process in place that reduces the probability and consequences of events that could negatively affect the 2020 Census programs ability to meet its objectives. The goal of the risk management process is to ensure a common, systematic, and repeatable assessment approach at both the program and project level so that risks can be effectively identified and managed, as well as clearly communicated to management, stakeholders, and executive level decision makers. Risk management is iterative and designed to be performed throughout the 2020 Census Program's research and testing, development, and implementation phases.

Specifically related to IT systems, the 2020 Census Program risk register includes risk mitigation strategies related to several risks:

- Technological innovations surfacing after the 2020 Census design is finalized;
- Cyber security incidents;
- Public perception of ability to safeguard response data;
- Funding requests not realized; and
- Late operational design changes.

Examples of mitigation strategies include:

- Build versatile operations and systems design;
- Establish a change control management process to assess impacts of change requests to facilitate decision-making;

- Prepare for rapid response to implement change based on the results of the change control process;
- Monitor system development efforts to ensure the proper security guidelines are followed;
- Audit systems and check logs to help in detecting and tracing an outside infiltration;
- Research other Census Bureau divisions, other government agencies, and the private sector to understand how they effectively mitigate the issue of public trust and IT security; and
- Prioritizing research, testing, and implementation activities per fiscal year to focus on those areas that can significantly impact cost and quality, and develop contingency plans to quickly respond to budget cuts.

Also, recognizing the need for timely decisions on which Census Enterprise Data Collection and Processing (CEDCaP) systems we need to build internally, and which we need to buy from external sources, we have engaged the Carnegie Mellon Software Engineering Institute to guide us and to ensure that the decision making process is successful. We created a technical design decision framework that determines key design capabilities based on business capabilities already identified in the 2020 Census Operational Plan, and helps conduct Capability Analysis and Assessment of in-house and Commercial-Off-The-Shelf (COTS) solutions.

We conducted the Fiscal Year 2016 System Readiness Baseline Assessment that was completed on January 12, 2016. It was then presented to, and approved by, the 2020 Census Program Executive Steering Committee on January 19, 2016. The changes to the 2016 Census Test business and capability requirements resulting from the assessment have been identified, and the IT system teams are working to complete the implementation of the baseline of business and capability requirements. For the Address Canvassing Test, the 2017 Census Test, and the 2018 End to End Test, the business and capability requirements also were reviewed and updated. Existing systems then will be assessed for their ability to fulfill the business and capability requirements.

The combination of the work that Carnegie Mellon is conducting, with the Capability Analysis and Assessment, will determine which systems will be built or bought, by March 2016. We will build out the chosen systems (COTS solutions need customization) and test them (whether they are built in-house, or COTS, or a combination) in the 2017 Census Test and 2018 Census End-to-End Test. Following the 2017 Census Test, we will revise systems as needed and use them in the 2018 Census End-to-End Test, implementing the planned census operations in real-time.

Use of Administrative Records

6. The Bureau's planned use of administrative records for the 2020 Census will rely on a large amount of data shared from one part of the federal government to another (the Bureau), most likely involving extremely large and repeated transfers of data across multiple systems.
 - Can you explain the specific steps the Bureau plans to take to secure the transfer of these

data as the Bureau expands its reliance on them to carry out the decennial census?

The Census Bureau has a long history of working with administrative data. The Census Bureau has safely received and used Federal tax information about both individuals and businesses in its non-decennial programs. It obtains Federal, state and commercial administrative data many times a year from a variety of sources for its ongoing research and statistical programs beyond the decennial census. In addition, the Census Bureau has some administrative records experience specifically in the decennial program. For example, since 1970, it has securely made use of administrative records to help enumerate group quarters, such as college dormitories and prisons. More recently, the Census Bureau conducted limited experiments during the Census 2000 and found there was potential to use administrative records to assist with follow-up and other operations, but that further research would be needed. Additionally, since 2000, the Census Bureau has used addresses provided by the U.S. Postal Service (USPS) Delivery Sequence File (DSF) as a starting point to update its Master Address File, a data file that contains a list of all known living quarters in the United States and Puerto Rico. For the 2010 Census, the Census Bureau used administrative records to help enumerate some group quarters and select cases for an operation that followed up on potentially inaccurate census responses. Throughout the decades, the Census Bureau has used secure processes for receiving and processing these data.

The Census Bureau has established data sharing agreements with key stakeholders, such as the Internal Revenue Service and the Social Security Administration. These formal agreements are developed in conjunction with the data provider and vetted by all required offices. This includes a legal review by both agencies to ensure that the transfer is secure, and that the data will be protected within the Census Bureau network. In addition, the IRS conducts Safeguard Reviews on site at the Census Bureau at regular intervals to assure that its data are being handled securely and in accordance with IRS laws, rules, and regulations.

Each agreement spells out the data transfer methods. These are mutually agreed upon by both agencies. All transfers of data are fully encrypted using approved NIST FIPS 140-2 encryption. The exact systems and how the data are to be used are also spelled out in the agreement. In addition, an Information Security Agreement is prepared, reviewed and signed by the appropriate management and designated points of contact at each agency.

Once at the Census Bureau, the data will reside on systems that are fully accredited according to NIST Special Publication 800-37r1 and assessed against NIST Special Publication 800-53r4 on a continuous basis.

- How can you be sure enumeration through imputation via administrative records will be as or more accurate than an in-person meeting with an enumerator?

The Census Bureau has extensive experience using administrative data to improve the accuracy, comprehensiveness and overall usefulness of many Census Bureau statistical programs, such as the Economic Census, the Longitudinal Household Dynamics Program,

the Small Area Income and Poverty Estimates program and more. These experiences give the Census Bureau familiarity with various administrative data files and with methods to assess accuracy and fitness for use. In fact, the Census Bureau appreciates the confidence in its capacity to acquire and properly use administrative data for a variety of statistical purposes, as expressed in the President's FY 2016 budget proposal, adopted by the Congress, to expand this expertise to serve the needs of a broader set of Federal statistical and evaluation offices.

Specific to the decennial census program, as outlined in the 2020 Census Operational Plan, after giving the population an opportunity to self-respond to the 2020 Census, the Census Bureau will use the most cost-effective strategy for contacting and counting people to ensure fair representation of every person in the United States.

Addresses for which the Census Bureau did not receive a self-response will form the initial universe of addresses for the Nonresponse Followup (NRFU) operation. Prior to any fieldwork, vacant addresses will be removed from the NRFU workload using administrative records. Undeliverable-As-Addressed information from the USPS will provide the primary administrative records source for the identification of vacant units. Vacancy status will be confirmed with at least one other administrative record or third-party data source. For example, if another source such as the Internal Revenue Service Individual Tax Returns File includes a return for an address identified as vacant by the USPS, it will not be removed from the NRFU workload.

After an initial attempt to contact nonresponding housing units, the NRFU workload will be further reduced through the removal of cases where administrative records and third-party data are available and usable to enumerate the occupied housing units. We will only do this when we believe we have good quality administrative records information for a particular address. If we don't, we'll make additional efforts to contact the household in person. The Examples of sources of administrative records and third-party data used to enumerate occupied housing units include: Internal Revenue Service Individual Tax Returns, Internal Revenue Service Information Returns, and Centers for Medicare & Medicaid Services Medicare Enrollment Database. The Census Bureau has decades of experience in acquiring and evaluating the quality of using these administrative records.

Addresses removed from the NRFU workload as either vacant or occupied will receive a final mailing that encourages occupants to self-respond to the 2020 Census. After each phase of the administrative records modeling, those addresses that are determined to be vacant will immediately be mailed a final letter encouraging self-response; for those addresses that are determined to be occupied and are incomplete after one personal visit attempt, a final letter encouraging self-response will be mailed after seven days. Research from the 2014 Census Test, 2015 Census Test, the 2010 Census, and the American Community Survey has helped us estimate that using administrative records and third-party data will allow for the following during the 2020 Census:

- *To remove approximately 11 percent of vacant housing units from the Nonresponse Followup universe; and*

- To enumerate approximately 16.5 percent of occupied nonresponding housing units after making one in-person visit.

The analysis of the 2015 Census Test demonstrates that when using administrative records to enumerate nonresponding housing units after one NFRU visit, we are able to capture sex and age data 100 percent of the time and race/Hispanic origin data 81.5 percent of the time. We will continue to evaluate these parameters as we progress with our research on the 2016 Census Test this fiscal year.

Additionally, the 2020 Census Operational Plan contains a detailed discussion (in Section 6.8) of how we are examining the quality impacts from the use of administrative records and third-party data. This research includes comparisons of these data to actual 2010 Census responses from matching households, and comparisons in our 2020 Census tests of the data quality from test panels that use administrative records to the data quality for panels that use traditional methods (additional personal visits followed by proxy reporting or imputation as needed). Where high-quality administrative records are not available from trusted sources, we will continue in-person visits to reach nonresponding housing units. Our examination of data quality also includes research regarding race and Hispanic origin data from administrative records and third-party sources.

To be clear, where high-quality administrative records are not available from trusted sources, we will continue in-person visits to reach nonresponding housing units.

- What specific steps is the Bureau taking to ensure their use will not have unintended consequences on the quality of the census?

As outlined in the 2020 Census Operational Plan, after giving the population an opportunity to self-respond to the 2020 Census, the Census Bureau will use the most cost-effective strategy for contacting and counting people to ensure fair representation of every person in the United States.

Where high-quality administrative records are not available from trusted sources, we will continue in-person visits to reach nonresponding housing units. Our examination of data quality also includes research regarding race and Hispanic origin data from administrative records and third-party sources.

Research from the 2014 Census Test, 2015 Census Test, the 2010 Census, and the American Community Survey has helped us estimate that using administrative records and third-party data will allow for the following during the 2020 Census:

- To remove approximately 11 percent of vacant housing units from the Nonresponse Followup universe; and
- To enumerate approximately 16.5 percent of occupied nonresponding housing units after making one in-person visit.

The analysis of the 2015 Census Test demonstrates that when using administrative records

to enumerate nonresponding housing units after one NFRU visit, we are able to capture sex and age data 100 percent of the time and race/Hispanic origin data 81.5 percent of the time. Research will continue as we progress with the 2016 Census Test this fiscal year.

Additionally, the 2020 Census Operational Plan contains a detailed discussion (in Section 6.8) of how we are examining the quality impacts from the use of administrative records and third-party data. This research includes comparisons of these data to actual 2010 Census responses from matching households, and comparisons in our 2020 Census tests of the data quality from test panels that use administrative records to the data quality for panels that use traditional methods (additional personal visits followed by proxy reporting or imputation as needed).

To be clear, where high-quality administrative records are not available from trusted sources, we will continue in-person visits to reach nonresponding housing units.

7. The Bureau seeks to use additional sources it currently does not have access to, such as NDNH and KidLink. However, the Bureau needs to make a decision soon on whether or not to use these sources. When will these decisions be made? Will this help lower costs?

As documented in the 2020 Census Operational Plan, the Census Bureau has already determined that the core set of administrative records data used to support the 2020 Census operations will include the following:

- *Internal Revenue Service Individual Tax Returns;*
- *Internal Revenue Service Information Returns;*
- *Centers for Medicare & Medicaid Services Medicare Enrollment Database;*
- *Indian Health Service Patient Database;*
- *Social Security Number Identification File;*
- *United States Postal Service Delivery Sequence File;*
- *United States Postal Service Undeliverable-As-Addressed Information;*
- *2010 Census Data; and*
- *American Community Survey Data.*

The Census Bureau will decide the final set of administrative records and third-party data to support the 2020 Census no later than September 2018. We believe the use of information from other sources, like the Supplemental Nutrition Assistance Program or the National Directory of New Hires (NDNH), will help improve the overall completeness and quality of data for households enumerated through the use of administrative records or third-party data.

Access to NDNH information, which would require a new statutory authority, would provide an important source of records for the 2020 Census, and further enable the Census Bureau to implement important cost-saving design changes. The NDNH data, in particular, would provide confirmatory information that would facilitate the use of federal tax information, which cannot under current agreement between the Internal Revenue

Service and the Census Bureau be used directly for enumeration. We believe that the value of accessing information in the NDNH and using it to corroborate data from the IRS, in conjunction with other administrative record and third-party data sources, would be approximately \$150 million (using 2010 figures and dollars). In their Progress Report on Administrative Records to Control Enumeration Costs, the GAO has indicated that this estimate is reasonable and recommends that we establish clear deadlines for accessing information from the NDNH.

The use of administrative records and third-party data will avoid over \$1.4 billion in costs during the 2020 Census (or \$9.80 per housing unit). For example, during the 2010 Census, the Census Bureau visited 14 million vacant housing units during Nonresponse Followup. By using information from the United States Postal Service, we will be able to remove approximately 11 percent, or 6 million vacant housing units from the workload during the 2020 Census, avoiding considerable costs.

8. In addition to NDNH and KidLink, there are nine additional sources the Bureau may use. When will the Bureau decide which database to use and again, what is the potential in lowering costs?

*As documented in the 2020 Census Operational Plan, the Census Bureau has already determined that the **core** set of administrative records data used to support the 2020 Census operations will include the following:*

- Internal Revenue Service Individual Tax Returns;
- Internal Revenue Service Information Returns;
- Centers for Medicare & Medicaid Services Medicare Enrollment Database;
- Indian Health Service Patient Database;
- Social Security Number Identification File;
- United States Postal Service Delivery Sequence File;
- United States Postal Service Undeliverable-As-Addressed Information;
- 2010 Census Data; and
- American Community Survey Data.

*The Census Bureau will decide the **final** set of administrative records and third-party data to support the 2020 Census no later than September 2018. We believe the use of information from other sources, like the Supplemental Nutrition Assistance Program or the National Directory of New Hires (NDNH), will help improve the overall completeness and quality of data for households enumerated through the use of administrative records or third-party data.*

Access to NDNH information, which would require new statutory authority would provide an important source of records for the 2020 Census, and further enable the Census Bureau to implement important cost-saving design changes. The NDNH data, in particular, would provide confirmatory information that would facilitate the use of federal tax information, which cannot under current agreement between the Internal Revenue Service and the

Census Bureau be used directly for enumeration. We believe that the value of accessing information in the NDNH and using it to corroborate data from the IRS, in conjunction with other administrative record and third-party data sources, would be approximately \$150 million (using 2010 figures and dollars). In their Progress Report on Administrative Records to Control Enumeration Costs, the GAO has indicated that this estimate is reasonable and recommends that we establish clear deadlines for accessing information from the NDNH.

The use of administrative records and third-party data will avoid over \$1.4 billion in costs during the 2020 Census (or \$9.80 per housing unit). For example, during the 2010 Census, the Census Bureau visited 14 million vacant housing units during Nonresponse Followup. By using information from the United States Postal Service, we will be able to remove approximately 11 percent, or 6 million vacant housing units from the workload during the 2020 Census, avoiding considerable costs.

Questions for The Honorable John H. Thompson

Director

U.S. Census Bureau

Questions from Representative Ted Lieu

Subcommittee on Government Operations

Hearing: "Preparing for the 2020 Census: Will the Technology be Ready?"

1. Please provide an update on what the Bureau is doing, both as part of the interagency working group, and on its own, to improve data collection on LGBT Americans.

The Census Bureau is a member of OMB's Interagency Working Group on Measuring Sexual Orientation and Gender Identity (SOGI). Through our participation in this group, we are learning from other Federal agencies about the current practices for collecting, editing, and reporting SOGI data. We are also gaining insight about the challenges and obstacles that agencies currently fielding SOGI questions have experienced, and benefiting from the technical solutions identified to address these issues. The lessons learned by our colleagues will be of critical importance should the Census Bureau be tasked with implementing SOGI questions in the decennial census or surveys, such as the American Community Survey (ACS).

As part of the Census Bureau's effort to research methods of enhancing our census and survey data using administrative records, staff have evaluated changes to individuals' first names and sex-coding in files from the Social Security Administration to identify people likely to be transgender. These data were then linked with 2010 Census data to evaluate demographic characteristics, residential patterns, and questionnaire response patterns.

The Census Bureau continues to improve the quality of its estimates of same-sex couples through testing and implementation of a new relationship question. The new question expands the husband or wife and unmarried partner categories to distinguish between same-sex and opposite-sex relationships. The new relationship question is now used in the American Housing Survey, the Survey of Income and Program Participation, and is being phased into the Current Population Survey. This question will be tested in the ACS in 2016 and is planned for implementation in the ACS in 2019, once the content testing cycle is completed. We also continue to test and work toward inclusion of the new question in the 2020 Census.

2. Please provide an update on what the Bureau is doing to improve data collection on gender identity, and transgender Americans, in particular.

The Census Bureau has collaborated with the National Center for Health Statistics (NCHS) on the testing and implementation of questions on sexual orientation in the National Health Interview Survey (NHIS). The NHIS is sponsored by NCHS and administered by the Census Bureau. From 2011 to 2012, the Census Bureau administered field tests to determine the best mode for collecting information on sexual orientation due

to the sensitivity of the questions. The questions on sexual orientation were then fully implemented in the NHIS in 2013 and, to date, remain part of the survey. The measures of sexual orientation used in the NHIS are now used across several Federal agencies spanning several Federal departments.

We are also collaborating with the Bureau of Justice Statistics (BJS) in the cognitive testing and planned implementation of SOGI questions in the National Crime Victimization Survey (NCVS). The NCVS is sponsored by BJS and administered by the Census Bureau. BJS has proposed the introduction of questions on sexual orientation and gender identity in the NCVS in July 2016. Cognitive testing of these questions has recently concluded, and survey instruments are being finalized.

As described above, the Census Bureau is harnessing administrative records to identify potential transgender Americans using changes to individuals' first names and sex-coding in files from the Social Security Administration.

3. Census Bureau data on same-sex couples only represents a subset of LGB people (since not all LGB people are in couples). That is why we need a direct measure of sexual orientation too: What specifically is the Census Bureau doing to resolve any methodological issues it sees as obstacles to adding sexual orientation and gender identity measures to the American Community Survey and other surveys.

The Census Bureau is committed to reflecting the information needs of our changing society and is constantly examining the effectiveness of census and survey data to collect accurate information on people and families. Gathering SOGI data more broadly will provide policymakers with a more accurate and comprehensive view of how LGBT Americans are faring in the United States today and how these trends change over time. Such information about the LGBT community's experiences with economic and social programs will prove useful not only to Federal policymakers but to all persons—academics, local decision-makers, and the general public-- concerned with the well-being of LGBT Americans.

At the same time, it must be understood that there are challenges to collecting SOGI data, including social desirability bias and interviewer reticence. But these challenges are not new; Federal statistical agencies have a long history of working together to leverage their expertise and resources to the benefit of the American public. As described above, the Census Bureau is actively engaged in the OMB Interagency Working Group (IWG) on Measuring Sexual Orientation and Gender Identity. While the work of the IWG is not yet complete, it is clear that there are a number of methodological issues that will need to be resolved before improved and wider measurement can occur. For example, there is little research on the effect of proxy reporting has on the accuracy of SOGI measurement. Proxy reporting is the practice of asking one member of the household to respond on behalf of other members in the household—an established way of significantly reducing data collection costs and burden. If proxy reporting yields inaccurate SOGI data, or a policy determination is made that SOGI questions cannot be asked through proxy reporting, that will affect our ability to measure SOGI more widely in established Federal surveys and censuses. The IWG is examining this issue now, and

testing is being planned. The Census Bureau does not have plans to test questions about gender identity or sexual orientation for the 2020 Census or the American Community Survey (ACS). However, as is the case to ensure the continued relevance of all the survey and census data collected by Census, we will work with OMB and other federal agencies to examine the changing requirements and data recommended for program implementation and use established best statistical practice as our guide.

Questions for Mr. Steven I. Cooper
 Chief Information Officer
 U.S. Department of Commerce

Questions from Chairman Mark Meadows
 Subcommittee on Government Operations

Hearing: "Preparing for the 2020 Census: Will the Technology be Ready?"

1. The Census Bureau released its first Operational Plan for the 2020 Census. Can you describe some specific steps the Bureau plans to take from here on out to improve the Bureau's scheduling practices and to meet the 2018 "dress rehearsal" deadline?

The 2020 Census Operational Plan outlines 350 major design decisions for the next decennial census, 168 of which have already been made. There are 182 decisions that still need to be decided, 167 prior to the 2018 Census End-to-End Test. The decisions will be data-driven, based on information collected from upcoming tests and operations. The dates for making these decisions are included in both the operational plan and the 2020 Census Lifecycle Schedule. More details about each of these decisions can be found in the Operational Plan at this link: <http://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/planning-docs/operational-plan.html>. The dates were determined to ensure adequate system testing and time to prepare for the 2018 End-to-End Test. Additionally, by the end of March 2016, the Census Bureau will baseline the integrated schedule for the 2018 Census End-to-End Test.

The Census Bureau has a number of program management efforts underway for the 2020 Census program that also will help ensure readiness for the 2018 Census End-to-End Test. For example, the 2020 Census Program has categorized the work of the 2020 Census in alignment with the enterprise work breakdown structure that is currently being developed by the Census Bureau. Scope, budget, and schedule are managed using this structure. To support the 2020 Census schedule development and maintenance, the Decennial Census Management Division has an office devoted to this work. Working closely with 2020 Census Program project teams, the office staff creates the schedule, routinely statuses activities on the schedule, evaluates the impact of change requests on the schedule, and monitors progress on a regular basis. The schedule includes both high-level items that plan the handoffs between the tests that precede the 2018 End-to-End Test and lower level items that plan requirements and system development tasks to ensure the ability to adequately test software before being used in each census test.

Finally, within the formal program and project-level risk management process, the Census Bureau develops mitigation plans for schedule risks, and contingency plans where appropriate. They are currently tracking risks related to the possibility of delays, late design changes, and major budget shortfalls.

2. Ms. Cha laid out several key IT decisions that have been deferred which may impact the Bureau's ability to have systems in place to implement the 2020 Census redesign. What

steps should the Bureau take to off-set this risk?

The Census Bureau recognizes the need to plan our remaining software development using a timely and comprehensive approach. They have a schedule of when critical IT decisions need to be made for the 2020 Census, which has been integrated with the CEDCAP schedule to assure that the critical functions needed to support the 2020 Census are delivered on time.

To off-set risk, they have engaged the Carnegie Mellon Software Engineering Institute to guide and to ensure that the decision making process is successful. They created a technical design decision framework that determines key design capabilities based on business capabilities already identified in the 2020 Census Operational Plan, and helps conduct Capability Analysis and Assessment of in-house and Commercial-Off-The-Shelf (COTS) solutions.

We conducted the Fiscal Year 2016 System Readiness Baseline Assessment that was completed on January 12, 2016. It was then presented to, and approved by, the 2020 Census Program Executive Steering Committee on January 19, 2016. The changes to the 2016 Census Test business and capability requirements resulting from the assessment have been identified, and the IT system teams are working to complete the implementation of the baseline of business and capability requirements. For the Address Canvassing Test, the 2017 Census Test, and the 2018 End to End Test, the business and capability requirements also were reviewed and updated. Existing systems then will be assessed for their ability to fulfill the business and capability requirements. The combination of the work that Carnegie Mellon is conducting, with the Capability Analysis and Assessment, will determine which systems will be built or bought, by March 2016.

Also, the Census Bureau has a formal, rigorous, and active risk management process in place at the program and project level, which takes into account both schedule and budget. This includes developing mitigation plans for all risks, and contingency plans where appropriate. This includes mitigation plans for several risks that relate to such things as the possibility of delays, late design changes, and major budget shortfalls.

3. Ms. Cha testified that one of the key deferred decisions is whether the Bureau plans to build or buy the needed IT solutions-which the Bureau does not plan to make until June 2016. Given there is now 14 months before preparations for end-to-end testing begins, is there enough time to build, test, integrate, scale, and secure systems in a 14 month period? In other words, is building a realistic option at this point?

The Census Bureau has accelerated the date for making these decisions to March 2016.

The Census Bureau is making innovative use of existing technology and software instead of inventing their own. They developed a working prototype for the 2020 Census that was successfully tested in 2015. Based on this prototype, they have incorporated most of the 2020 Census Program specifications into the next iteration. They will finish this work

based on the results of tests in 2016 and 2017 in time for the 2018 End-to-End Test.

The Census Bureau has engaged the Carnegie Mellon Software Engineering Institute to guide and to ensure that the build-or-buy decision-making process is successful. They created a technical design decision framework that determines key design capabilities based on business capabilities already identified in the 2020 Census Operational Plan, and helps conduct Capability Analysis and Assessment of in-house and Commercial Off-The-Shelf (COTS) solutions.

We conducted the Fiscal Year 2016 System Readiness Baseline Assessment that was completed on January 12, 2016. It was then presented to, and approved by, the 2020 Census Program Executive Steering Committee on January 19, 2016. The changes to the 2016 Census Test business and capability requirements resulting from the assessment have been identified, and the IT system teams are working to complete the implementation of the baseline of business and capability requirements. For the Address Canvassing Test, the 2017 Census Test, and the 2018 End to End Test, the business and capability requirements also were reviewed and updated. Existing systems then will be assessed for their ability to fulfill the business and capability requirements. The combination of the work that Carnegie Mellon is conducting and the assessment of system readiness of in-house system development will determine which systems will be built or bought.

4. In your capacity as the Department-level CIO, which has been recently heightened by FITARA, please describe the ways in which you have been involved with the CEDCAP program in order to mitigate the risk of being without a permanent CIO.

Your Committee recently published the first Federal IT Acquisition Reform Act (FITARA) Implementation Scorecard, upon which the Department of Commerce (DOC) received the best overall grade in the Federal Government.

The scorecard reflects the Department's progress in successfully establishing and implementing data center consolidation plans, IT portfolio review and savings, risk assessment transparency, and the maturation of CIO authorities. The primary reason we have been able to achieve these successes is our strategic alignment and adherence with previous Government IT initiatives like the 25-Point Implementation Plan to Reform Federal Information Technology (IT) Management; Office of Management and Budget (OMB) Memorandum M-11-29, Chief Information Officer Authorities; and OMB Memorandum M-12-10, Implementing PortfolioStat.

Common themes of all of these directives are IT transformation, improvement of operating efficiencies, leveraging shared services, and increasing Chief Information Officer (CIO) accountability and responsibility for all IT spend. FITARA further clarifies a CIO's role, accountability, and relationships. FITARA also provides the tools that I can use to further mature the governance and oversight that we as an agency have already put in place.

The implementation of key information technology management processes including an enterprise systems development lifecycle and program lifecycle used to govern the Census Enterprise Data Collection and Processing (CEDCaP) program have given myself, other managers, and oversight bodies the tools and data required to ensure alignment with our goals, priorities, and compliance with statutory requirements. They have also provided a timely, consistent, and transparent source for reliable cost, schedule, and performance status. In addition, I regularly engage with the Acting CIO of the U.S. Census Bureau, who is the accountable Associate Director for delivering the CEDCaP solution. The Acting CIO is also part of the DOC CIO Council, the body that is responsible for DOC's collective compliance with other memorandums and directives.

CEDCaP is listed as a critical project as part of the Department of Commerce's Acquisition Framework and is a prime example of the successes the Department of Commerce and U.S. Census Bureau have collaboratively achieved in the adherence, alignment, and implementation of the objectives of these directives.

Through coordinated efforts including my own oversight via the aforementioned processes and the Department's internal Information Technology Review Board, the CEDCaP program achieved the DOC Milestone 2 Review Approval from the Deputy Secretary of Commerce in July 2015, and also has accomplished the following:

- *Product Release 1 - CAPI Tech Refresh complete;*
- *CEDCaP Capabilities released to support CAPI Tech Refresh (ACS & DEMO);*
- *Questionnaire Design and Metadata (Content Metadata - CoMet) release;*
- *Address Listing and Mapping (LiMA) release;*
- *Survey (and Listing) Interview Operational Control (Mobile Case Management - MCM) release;*
- *Dashboard for Monitoring (Unified Tracking System- UTS) release;*
- *CoMet released to production in support of 2015 COS/ASM;*
- *Service Oriented Architecture (SOA) released to production;*
- *Finalized Program-Level Systems Engineering Management Plan (SEMP) & Test and Evaluation Master Plan (TEMP);*
- *Finalized Segment Architecture and Transition Plans (second iterations);*
- *Developed CEDCaP Program-level Test and Evaluation Management Plan (TEMP), CEDCaP Program-level Systems Engineering Management Plan (SEMP) and CEDCaP Program level End-to-End (e2e) Integration Testing Framework and Operational Model;*
- *Implemented an enterprise IT Acquisition for Business Services Support (ITBSS) Enterprise Service Framework (ESF) with CEDCaP Task Order for program management support;*
- *Successfully conducted testing of Centurion and an alternate application – Primus, in a cloud-based environment;*
- *Implemented an enterprise IT Acquisition for Systems Engineering and Integration (SE&I) Services contract;*
- *Developed Request for Quote (RFQ) for Commercial off-the-shelf (COTS)*

- product(s) for assessment prior to final build/buy decisions;
- Performing COTS assessment on the analysis of alternatives that includes comparing in-house custom solution(s) versus COTS product(s); and
- Engaged the Carnegie Mellon (CM) Software Engineering Institute (SEI) to perform an independent assessment of CEDCaP's COTS assessment approach (e.g. inputs, evaluation criteria, process).

We are in the process of selecting an Independent Validation and Verification (IV&V) contractor for CEDCaP to further mitigate risk and ensure the program stays on track. We expect to complete our selection by the second quarter of 2016.

5. What can the 2020 Census Directorate and the IT Directorate do to better ensure that the IT Directorate can successfully deliver the systems to support the 2020 Census?

The Census Bureau has gone a long way to improving the communication and integration of their respective Directorates to ensure the successful delivery of the 2020 Census. For example, the 2020 Census Directorate has a Chief of the Decennial IT Division (DITD) that oversees the 2020 Census System Engineering and Integration (SE&I) team. This person and the Associate Director of Decennial Census Programs meets with the Census Bureau's Chief Information Officer (CIO) on a bi-weekly basis to ensure that the objectives for the 2020 Census and associated tests are aligned with enterprise IT plans and that tasks are prioritized based on meeting those objectives.

Further, the Chief of DITD works with the IT Directorate under the CIO to ensure that the solutions relating to Decennial IT are following the Census Bureau's IT policies and procedures, and that the processes for their development are in alignment with the Enterprise Architecture, enterprise Systems Development Life Cycle (eSDLC), security, and standards.

While the Chief of DITD is responsible for providing systems for the 2020 Census, the hardware and infrastructure (telecommunications, LAN/WAN, data center, backup/recovery, mobile devices, Cloud, security, etc.) are provided by the IT Directorate under the CIO. Therefore, there is shared responsibility and accountability. All systems managed by the DITD are assessed and approved by the Office of Information Security (OIS), independent of DITD. No system that DITD builds will go into production until the CIO signs off on the Authority-To-Operate (ATO). Independent assessment and approval ensures consistent approach to security. The IT Directorate has also hired qualified Security Engineers and embedded them into the 2020 Program to ensure security is built into every system or solution from the beginning of every project.

The IT Directorate established the enterprise systems readiness process being used for the 2020 Census Directorate to assess and ensure system readiness, and has embedded a Solution Architect, Chief Program Engineer, and Systems Engineers as dedicated members of the 2020 Census SE&I Team.

In addition, the Census Bureau has conducted multiple sessions where the leaders of the

2020 Decennial Census and IT directorates were brought together to discuss and roll out plans, schedules and an IT roadmap for the 2020 Census, which were signed off by both directorates.

Decennial IT and the IT Directorate collaborate frequently and are in-sync from planning through execution regarding the 2020 Census. They also jointly attend the Executive Steering Committees at the Bureau and program levels.

Overall, the IT and Decennial leaders understand the complexity, importance, and critical nature of the 2020 Census and are working together embracing a shared leadership model. I feel confident that the continuation and maturation of these processes between the two Directorates will provide the foundation for the successful delivery of IT solutions for the 2020 Census.

6. The Bureau maintains the quality of administrative records data by ensuring its accuracy, relevance, and timeliness, for their planned uses. Do you believe the Bureau has been taking the right steps? What else could the Bureau do in order to maintain the quality of the records?

The Census Bureau is committed to designing the most cost-effective strategy for contacting and counting people to ensure fair representation of every person in the United States.

The Census Bureau has a long history of working with administrative data, as far back as 1970, to help with the enumeration of group quarters, colleges, and prisons. The 2020 Census Operational Plan contains a detailed discussion of how the Bureau is examining the quality impacts from the use of administrative records and third-party data. As documented in the operational plan, the Census Bureau has already determined that a core set of high quality administrative records data will be used to support the 2020 Census operation, including:

- *Internal Revenue Service Individual Tax Returns;*
- *Internal Revenue Service Information Returns;*
- *Centers for Medicare & Medicaid Services Medicare Enrollment Database;*
- *Indian Health Service Patient Database;*
- *Social Security Number Identification File;*
- *United States Postal Service Delivery Sequence File;*
- *United States Postal Service Undeliverable-As-Addressed Information;*
- *2010 Census Data; and*
- *American Community Survey Data.*

Census Bureau research includes comparisons of these data to actual 2010 Census responses from matching households. Research also compares data quality across 2020 Census Test panels using administrative records and traditional methods. Where high-quality administrative records are not available from trusted sources, the Census Bureau will continue in-person visits to reach nonresponding housing units.

I believe these are the right steps to take to ensure the appropriate use of administrative records in conducting an accurate and inclusive 2020 Census.

7. The Bureau seeks to use additional sources it currently does not have access to, such as NDNH and KidLink. However, the Bureau needs to make a decision soon on whether or not to use these sources. When will these decisions be made? Will this help lower costs?

The Census Bureau will decide the final set of administrative records and third-party data to support the 2020 Census no later than September 2018. We believe the use of information from sources like NDNH and KidLink will help improve the overall completeness and quality of data for households enumerated through the use of administrative records or third-party data.

Access to NDNH information, which would require a new statutory authority, would provide an important source of records for the 2020 Census. In particular, the data would provide a critical confirmation link to support the use of federal tax information, which cannot under current agreement between the Internal Revenue Service and the Census Bureau be used directly for enumeration.

We believe that the value of accessing information from the NDNH and using it to corroborate data from the IRS, in conjunction with other administrative record and third-party data sources, would be approximately \$150 million (using 2010 figures and dollars). In their Progress Report on Administrative Records to Control Enumeration Costs, the GAO has indicated that this estimate is reasonable and recommends that the Census Bureau establish clear deadlines for accessing information from the NDNH.

Using administrative records will avoid over \$1.4 billion in non-response follow-up costs during the 2020 Census.

8. In addition to NDNH and KidLink, there are nine additional sources the Bureau may use. When will the Bureau decide which database to use and again, what is the potential in lowering costs?

As documented in the 2020 Census Operational Plan, the Census Bureau has already determined that the core set of administrative records data that will support the 2020 Census operation will include:

- *Internal Revenue Service Individual Tax Returns;*
- *Internal Revenue Service Information Returns;*
- *Centers for Medicare & Medicaid Services Medicare Enrollment Database;*
- *Indian Health Service Patient Database;*
- *Social Security Number Identification File;*
- *United States Postal Service Delivery Sequence File;*
- *United States Postal Service Undeliverable-As-Addressed Information;*
- *2010 Census Data; and*
- *American Community Survey Data.*

The Census Bureau will decide the final set of administrative records and third-party data to support the 2020 Census no later than September 2018.

Using administrative records will avoid over \$1.4 billion in non-response follow-up costs during the 2020 Census. In their Progress Report on Administrative Records to Control Enumeration Costs, the GAO has indicated that this estimate is reasonable and recommends that the Census Bureau establish clear deadlines for assessing information from the NDNH.

Enclosure

Chairman Will Hurd
Chairman Mark Meadows
Written Questions for the Record to
Mr. Robert Goldenkoff
“Preparing for the 2020 Census: Will the Technology be Ready?”

The Bureau's Upcoming and Important Deadlines

- 1. Do you believe the Bureau will meet all required deadlines for the 2018 end-to-end test, also known as "dress rehearsal" for the 2020 Census? Why or why not?**

Bureau officials have stated that final decisions on the use of administrative records are needed by the end of fiscal year 2017 in order for those records to be included in the Bureau's 2018 end-to-end test. As noted in our statement, the Bureau has already committed to three uses of administrative records, and has identified nine potential additional uses of administrative records that may help improve the quality of the census still further. While the Bureau is researching the feasibility of these additional uses, as of August 2015, the Bureau had not set deadlines for making these decisions, nor had it set deadlines for determining the specific administrative records that will support those uses. Deadlines for when final go/no-go decisions need to be made will help ensure the records are adequately tested during the end-to-end test. As such, in our October 2015 report,¹ we recommended that the Bureau establish clearly documented deadlines for making final decisions about which records to use, and for what purposes. The Bureau agreed and said it will prepare an action plan to document the steps it is taking to implement our recommendation. Moreover, GAO has been asked by the Committee on Oversight and Government Reform to assess the reliability of the Bureau's integrated schedule, and we expect to carry out that assessment in 2016.

- 2. In your opinion, how has the Bureau's delay in making key decisions affected the 2020 Census?**

We have not conducted the work necessary to answer this question. However, in 2013, the Bureau noted that the effects of the continuing resolution and sequestration, as well as the potential for future budget uncertainty later in the decade, necessitated a reassessment of priorities in the research and testing program. The Bureau delayed and consolidated previously planned tests to focus on cost cutting. The new testing focus could not supply information necessary for decision making by the end of the 2014 fiscal year as originally planned. As such, decisions on preliminary design were postponed from September 2014 to September 2015. Moreover, in our prior work, we have underscored the importance of staying on schedule so that key activities can occur at

¹GAO, *2020 Census: Additional Actions Would Help the Bureau Realize Potential Administrative Records Cost Savings*, GAO-16-48 (Washington, D.C.: Oct. 20, 2015).

the appropriate time and in the correct sequence in order to meet key dates for testing activities, as well as for collecting, analyzing, and disseminating information.

GAO has been asked by the Committee on Oversight and Government Reform to assess the reliability of the Bureau's integrated schedule, including what schedule risk analyses the Bureau may have conducted that might inform this question, and we expect to carry out that assessment in 2016.

Bureau's Use of Administrative Records

3. **The use of administrative records is not new for the Bureau, but the Bureau plans to rely on them more heavily than ever before. Moreover, while some uses have been decided, a number of other uses are still being debated.**

- a. **Are you concerned about this ongoing debate?**

As of August 2015, the Bureau had not set deadlines to determine when to make final decisions on which of its 12 identified uses of administrative records it will implement for the 2020 Census, nor had it set deadlines for determining exactly which records from which sources it will tap in support of each use it implements. Setting deadlines will help the Bureau complete needed activities on time and prioritize which activities—or records—to abandon if time and resources become a constraint. As with much of the design of the 2020 Census, most of the uses still being considered are new. Moving forward, to help support broader public discussion on trade-offs that the Bureau may need to make on the role of administrative records in the 2020 Census, the Bureau should address our prior recommendation to develop and implement an effective congressional outreach strategy, particularly on new design elements the Bureau is researching and considering, as well as on cost-quality trade-offs of potential design decisions.

- b. **What challenges does the Bureau face in ensuring that administrative records uses are successfully implemented?**

Our October 2015 report listed three key challenges the Bureau faces in ensuring that administrative records uses are successfully implemented:

- **Data Quality.** Although the Bureau has no control over the accuracy of data provided to it by other agencies, it is responsible for ensuring that data it uses for 2020 Census are of sufficient quality for their planned uses. Data quality can involve the accuracy, relevance, and timeliness of the data. The Bureau has taken steps to ensure the quality of the records it is considering using for 2020 and plans comprehensive testing of all records during an end-to-end test of its 2020 Census design (to be conducted in 2018). Additionally, the Bureau plans testing of administrative records for the 2016 Census Test in the Los Angeles and Houston metro areas, in a large test of address canvassing, also in 2016, and in an another site test in 2017 at a to-be-determined location.
- **Prevention of Disclosure.** We have previously reported that until the Bureau implements a complete and comprehensive security program, it will have limited

assurance that its information and systems are being adequately protected against unauthorized access, use, disclosure, modification, disruption, or loss. In response, the Bureau has cited examples of its past long-standing experience in collecting data from other agencies and reporting on it as evidence of the Bureau's ability to prevent disclosure of information. Once the Bureau obtains access to an administrative data source, it transfers the information that it needs to Bureau servers and maintains the information within the Bureau's firewalls and information security infrastructure.

- **Public Perception.** A third challenge is the extent to which the public will accept the sharing of personal data across government agencies for the purposes of the census. In 2013, the Bureau contracted for regular polling of nationally representative individuals on the extent to which they prefer data to come from information already provided to federal and state governments, or from a survey they fill out. Findings included that respondents were evenly divided when asked whether they prefer the Bureau to obtain someone's name and age directly from the Social Security Administration rather than asking for this information on a questionnaire. In 2013, the Bureau began hosting quarterly program management reviews encouraging dialogue with oversight on selected technical aspects of the Bureau's ongoing research and testing. These reviews are open to the public and viewable online. These supplement the Bureau's monthly status reports on ongoing research projects that the Bureau provides to Office of Management and Budget and, later, Congress. Additionally, the Bureau is developing a communications campaign for 2020, which it will formally launch in 2016. The campaign will include information about how the Bureau intends to use administrative records in the 2020 Census.

Given the many potential uses of administrative records the Bureau has identified, it will be important for the Bureau's external messaging to consider the range of uses. For example, some people may feel differently about the Bureau using administrative records for enumerating, as opposed to targeting the time of day they will be contacted by the Bureau. As noted above, the Bureau should address our prior recommendation to develop and implement an effective congressional outreach strategy.

4. The Bureau and GAO both appear to be in agreement that the expanded use of administrative records could save significant amounts of money compared to historical Census operations. Can you explain whether these "cost savings" are real savings to the taxpayer or just "paper savings"?

The Bureau's 2015 Census Test appears to have demonstrated the feasibility and potential effectiveness of administrative records for several uses during nonresponse follow-up, which the Bureau estimates could save up to \$1.4 billion compared to using traditional census methods in the 2020 Census. The Bureau recently estimated that with a new approach it can conduct the 2020 Census for \$12.3 billion, which is around \$5 billion less than the Bureau's initial cost estimate of \$17.5 billion to repeat the design and methods of the 2010 Census. The Bureau plans to do this in part by reengineering key census-taking methods, including making greater use of information already provided to federal and state governments as they administer programs—administrative records. We

have ongoing work to assess the reliability of the Bureau's 2020 Census lifecycle cost estimate. We plan to inventory the cost assumptions provided by the Bureau and identify those related to changes from historical census designs and with significant cost-saving potential.

5. In preparation for the 2016 and 2017 tests, does GAO recommend the Bureau to use administrative records differently than it did in its 2015 tests? If so, how?

GAO has not conducted work with which to recommend that the Bureau use administrative records differently than it did in 2015. However, according to Bureau officials, the results of the 2015 test demonstrated that it can save an estimated \$1.4 billion by using administrative records in the ways it already committed to and tested in 2015. Specifically, in 2015, the Bureau tested, among other things, how well administrative records; helped identify vacant addresses in advance of follow-up field work; substituted for additional visits to collect information from nonresponding households and from proxies, such as neighbors; and predicted the best times for sending enumerators to households to increase the likelihood that someone is home.

The Bureau plans to test additional uses of administrative records beyond what it did in 2015, such as using them for group quarters populations (e.g., dormitories, prisons, nursing homes, or homeless shelters), and for respondent validation for those responding without a unique identification number.

6. In your October 2015 report, GAO recommends the Bureau should clearly set deadlines for making decisions on the use of records it does not have access to, such as NDNH and KidLink.

a. Do you recommend the Bureau uses them? Why or why not?

GAO has not done work with which to recommend the use of specific administrative records for the 2020 Census. However, if accessible and used properly, these data sources have the potential to improve the quality of the census and potentially reduce its cost. Thus, going forward, it will be important for the Bureau to continue to explore the feasibility of using these records, ensuring, for example, that the records improve cost and quality, and that the Bureau understands the IT implications associated with expanded use of administrative records. At the same time, the Bureau has a lot of activity to follow through on in the time remaining, so it will be important to reduce the number of open options to ones the Bureau can manage well.

The National Directory of New Hires (NDNH) is a national database of wage and employment information used for child support enforcement. Bureau officials believe that name and wage information from NDNH could help corroborate the tax data from IRS that the Bureau already has access to, improving the collective accuracy of the records. The President's 2016 budget submission included a request for legislation that would authorize the Department of Health and Human Services to share NDNH data with the Bureau for statistical purposes such as the decennial census. KidLink is a database from the Social Security Administration (SSA) that

links parent and child Social Security numbers for children born after 1998 in U.S. hospitals. It is valuable to the Bureau, because children, and babies less than 1 year old in particular, have been historically undercounted. Bureau officials have said that this database could help identify another 1 million people. According to the Bureau, SSA raised issues about Bureau access to these data. However, the Bureau maintains that it has access authority for this data. The Census Bureau Director says that he will work with departmental staff and the Office of Management and Budget to explore an administrative solution that would provide the Bureau with access. Bureau officials said that there is value in accessing these records for the Bureau's other statistical surveys as well.

b. How do you suggest the Bureau go about making this decision?

Bureau officials have said that they are examining ways to quantify the potential effect that their access to these additional sources could have on the 2020 Census. We believe this step is critical to the decision making process. The Bureau estimated that the value of acquiring NDNH and using it to corroborate data from IRS, in conjunction with other administrative and third-party data sources, would be approximately \$157.5 million (using 2010 figures and dollars). This assumes no nonresponse follow-up visits for cases with administrative and/or third-party data, so that the actual savings would likely be less since the Bureau recently decided to make at least one follow-up visit before enumerating a household with administrative records.

c. What are your suggested deadlines?

The Bureau will need to consider when to end research pursuits that show less promise for substantially reducing the census' cost or meeting other 2020 goals so that it can focus resources on successfully refining and implementing activities that have greater potential. Knowing deadlines for when final go/no-go decisions need to be made about which records the Bureau will use, how it will use them, and for which purposes will help ensure necessary activities are completed on time. Deadlines regarding still uncertain purposes or those involving records the Bureau is still pursuing, such as NDNH and KidLink, as well as those from some states, will also help the Bureau prioritize which activities—or records—to continue pursuing or to abandon if time becomes a constraint. Bureau officials state that there is value in accessing these records for the Bureau's other statistical surveys as well, and that even if they are unable to obtain the additional records in time for the 2020 Census, they would continue pursuing them for these other purposes, as well as for use in future censuses.

Enclosure

**Post-Hearing Questions for the Record
Submitted to Carol Cha
From Chairman Will Hurd and Chairman Mark Meadows
“Preparing for the 2020 Census: Will the Technology Be Ready?”**

The Bureau's Upcoming and Important Deadlines

1. Given that less than 2 years remain before preparations for the end-to-end testing begins, what specific steps should the Bureau take to ensure CEDCAP systems are ready in time?

The Bureau should take several key steps to better ensure the readiness of CEDCAP systems in time for planned end-to-end testing:

- Decide as soon as possible whether to build or buy each of the needed CEDCAP systems and establish risk mitigation plans to address the unique challenges associated with each of these approaches. For example, the build approach will require a significant amount of technical expertise and time to build, integrate, and test unproven system capabilities. The buy approach, as we have previously reported, can also require a significant amount of time (e.g., in preparation for the 2010 Decennial Census, the Bureau was issuing solicitations and awarding contracts in 2005).¹
- Promptly address open GAO recommendations related to 2020 IT planning and information security. Consistent with our IT planning recommendations, identify the deferred decisions with the largest impact on the CEDCAP systems (e.g., IT infrastructure and scalability needs) and make those decisions as soon as possible. For those deferred decisions impacting the development and/or acquisitions of IT systems, but which the Bureau does not believe it can make soon, consider eliminating them from the scope of the 2020 Census design.
- Hire a permanent Chief Information Officer for the Census Bureau as soon as possible. Until this critical vacancy is filled, the Commerce Chief Information Officer should take a greater role in overseeing the CEDCAP program to help ensure systems will be ready in time.
- Take steps to ensure that the CEDCAP implementation schedule is driven by the 2020 Census program. As I previously testified, the CEDCAP program is intended to deliver an enterprise-wide system-of-systems to be used for all of the Bureau's survey data collection and processing functions, including those used by the American Community Survey and Economic Census.² While CEDCAP is releasing capabilities for these surveys, the 2020 Census must take priority in light of the immovable deadline and expected reliance on these systems to achieve its projected \$5.2 billion in savings.

¹GAO, *Information Technology: Agencies Need to Establish and Implement Incremental Development Policies*, GAO-14-361 (Washington, D.C.: May 1, 2014).

²GAO, *2020 Census: Key Information Technology Decisions Must Be Made Soon*, GAO-16-205T (Washington, D.C.: Nov. 3, 2015).

2. Based on GAO's experience with system testing, what would be the effect if the Bureau is unable to have the systems in place in time for end-to-end testing in a census-like environment?

The purpose of Bureau's planned end-to-end test is to ensure that all the various operations and systems are integrated before the decennial census. If all systems are not in place in time for this critical test, there is greater risk that they will not operate as intended during the 2020 Census.

As we previously reported, the 2010 Census experienced significant operational issues largely because the end-to-end testing was not comprehensive. Although multiple GAO reports, testimonies, and recommendations urged the Bureau to conduct full end-to-end testing leading up to the 2010 Census, the Bureau removed several key operations from this test. One such system excluded from the test was the paper-based operational control system. During the 2010 census-taking operations, this system experienced outages, slow performance, and problems generating and maintaining timely reports. As a workaround, the Bureau had to severely limit the number of concurrent users on the system.³

Another example of a large-scale system roll-out that experienced significant issues due to incomplete testing is Healthcare.gov. Specifically, we reported that Healthcare.gov lacked complete integration and end-to-end testing of the system, and as a result the Department of Health and Human Services' Centers for Medicare & Medicaid Services lacked a basis for knowing if all Healthcare.gov interconnected systems could operate correctly, pass data correctly to one another, and store data correctly prior to system launch.⁴ This was a key contributor to the problems Healthcare.gov faced, including inadequate system capacity, numerous errors in software code, and limited system functionality. Consequently, consumers attempting to enroll in health plans were met with confusing error messages, slow load times for forms and pages, and, in some cases, website outages.

For the 2020 Census, the Bureau's redesigned operations are expected to be more heavily dependent upon IT systems than ever before. Accordingly, comprehensive end-to-end testing that includes these new procedures and system interfaces will be critical to a cost-effective and complete census.

3. Is the Bureau's planned August 2017 milestone to begin preparations for the end-to-end system testing a reasonable date?

Yes, consistent with my remarks at the hearing, the planned August 2017 milestone date is reasonable and would likely provide the Bureau with sufficient time to correct defects identified during testing before the 2020 Census—which the Bureau currently plans to address from July 2018 to May 2019.

However, it is important to note that this time frame for achieving end-to-end testing is only feasible if the Bureau rapidly moves away from researching and testing prototype systems, as I previously testified, and instead focuses on implementing and testing the production

³GAO, *2010 Census: Data Collection Operations Were Generally Completed as Planned, but Long-standing Challenges Suggest Need for Fundamental Reforms*, GAO-11-193 (Washington, D.C.: Dec. 14, 2010), and *Information Technology: Census Bureau Testing of 2010 Decennial Systems Can Be Strengthened*, GAO-09-262 (Washington, D.C.: Mar. 5, 2009).

⁴GAO, *Healthcare.gov: CMS Has Taken Steps to Address Problems, but Needs to Further Implement Systems Development Best Practices*, GAO-15-238 (Washington, D.C.: Mar. 4, 2015).

systems that will be used for 2020 operations. Specifically, while the Bureau has conducted several tests to date, these tests have been using prototype systems in order to define business requirements—and the Bureau is continuing to refine these prototype systems for the next major site test in the spring of 2016. However, given the limited time remaining, continuing to test and refine prototype systems rather than production systems will result in diminishing returns. The Bureau should take the steps described in my answer to question 1 to ensure that the production systems are ready in time for end-to-end testing.

4. What additional security and IT infrastructure measures should the Bureau take in its 2016 and 2017 tests before the 2018 end-to-end test?

As indicated in my prior response, in 2016 and 2017 the Bureau should focus on acquiring, implementing, securing, and testing production systems, rather than testing and refining prototype systems that will not be used in the 2020 Census. To do so, the Bureau must make the decisions about the projected demand that the IT infrastructure would need to meet and whether it will build or buy the needed production systems soon. Until these decisions have been made, the Bureau cannot effectively identify all appropriate security and IT infrastructure measures that it must take to support 2020 operations (and the testing leading up to these operations).

5. In your statement, you explain some key IT decisions that must be made quickly.

○ **How long have these decisions been deferred?**

The Bureau has deferred key IT decisions for between roughly 2 and 4 years. Specifically, the Bureau had originally planned to complete all research and testing activities and determine the preliminary design decisions for the 2020 Census by September 2014. Many decisions about the redesign were documented in last month's 2020 Census Operational Plan; however, others were deferred because selected research and testing activities are still ongoing. As detailed in my written testimony, the Bureau deferred many key IT-related decisions that will impact the development of the operational system to 2016 through 2018.

○ **What are some repercussions in the deference?**

By making key IT-related design decisions too late, the Bureau may not be able to fully implement one or more of its planned redesign areas, such as optimizing self-response (via the Internet response option) and reengineering field operations (by automating and optimizing enumerator case assignment and routing and by using mobile devices for data collection). Since the planned systems are large and complex, the Bureau could simply run out of time to put them in place, thus diminishing the projected \$5.2 billion in cost savings.

○ **What are some foreseeable long-term effects?**

A significant long-term effect is the possibility of perpetuating the Bureau's cycle of building and/or acquiring unique, survey-specific systems with redundant capabilities for each survey it delivers. As reported by the Bureau, over \$1 billion was spent to build and deploy systems that supported the 2010 Census, which were retired following that decennial.⁵ To the Bureau's credit, its CEDCAP program was

⁵The Office of Management and Budget's IT Dashboard.

established to break that inefficient cycle by delivering an enterprise-wide system-of-systems to be used for all of the Bureau's survey data collection and processing functions. However, if CEDCAP systems are not ready in time to support the 2020 Census, the Decennial Census program may revert back to developing or acquiring separate systems to serve its needs.

As a result of the Bureau's challenges in key IT internal controls and its looming deadline, we identified CEDCAP as an IT investment in need of attention in our February 2015 high-risk report.⁶ We recently initiated a review of the CEDCAP program for your subcommittees, and expect to issue a report in the spring of 2016.

- o **What is the first and foremost important decision that has to be made, and why?**

The first and most important decision the Bureau needs to make, which will enable it to begin IT implementation, is whether it will build or buy the IT solutions needed to support the 2020 Census redesign, such as the operational control system and Internet and mobile data collection systems, for the reasons outlined in my response to question 1.

- 6. A number of key leadership positions at the Bureau remain vacant, including the positions of Chief Information Officer and Assistant Chief of Business Integration for the Bureau's CEDCAP initiative. How have these vacancies affected the Bureau and its planning for the 2020 Census?**

As of November 2015, one of the key IT leadership positions has been filled, according to Bureau officials (CEDCAP Assistant Chief of Business Integration); however, a new vacancy will be opening soon (CEDCAP Program Manager). According to Bureau officials, they are currently seeking a replacement, in addition to their other ongoing efforts to fill the remaining leadership vacancies.

It will be important for these vacancies to be filled as soon as possible in order to improve the Bureau's ability to effectively deliver CEDCAP systems in time for 2020 operations. For example, with the newly filled Assistant Chief of Business Integration position, the Bureau will be better positioned to oversee the integration of schedule, risks, and budget across the 12 CEDCAP projects. The remaining gaps, including the Chief Information Officer, must be closed to ensure that the CEDCAP program will be adequately governed and held accountable for cost, schedule, and performance plans.

- 7. In addition, the Bureau's IT staff is "highly-matrixed." Has this made project accountability and decision-making more difficult? How has it affected costs?**

We have ongoing work for your subcommittees that will include examining the Bureau's use of IT staff in both the CEDCAP and 2020 Census programs.

⁶Every 2 years at the start of a new Congress, GAO calls attention to agencies and program areas that are high risk due to their vulnerabilities to fraud, waste, abuse, and mismanagement, or are most in need of transformation. As part of a new entry into the February 2015 update to our High-Risk Series focused on improving the management of IT acquisitions and operations, CEDCAP was identified as an IT investment—among others across the federal government—in need of the most attention. See GAO, *High-Risk Series: An Update*, GAO-15-290 (Washington, D.C.: Feb. 11, 2015).

8. How is the Bureau's current IT posture and implementation of security options? Do you have any recommendations for improvement?

In January 2013, we reported that the Bureau had a number of weaknesses in information security controls intended to limit access to its systems and information, as well as those related to managing system configurations and unplanned events. We attributed these weaknesses to the fact that the Bureau had not fully implemented a comprehensive information security program, and made 115 recommendations aimed at addressing these deficiencies.⁷ As of October 29, 2015, the Bureau had addressed 66 of the 115 recommendations, and of the remaining open recommendations, we had determined that 30 required additional actions by the Bureau while the other 19 were still under evaluation. The Bureau's progress toward addressing our security recommendations is encouraging. However, more work remains to address the open recommendations. Additionally, we have ongoing work reviewing the CEDCAP program that will allow us to better understand the specific IT security issues related to the 2020 Census design and make associated recommendations where appropriate.

9. The Bureau seeks to generate the largest possible self-response in the history of the decennial census in order to reduce the number of households requiring follow-up.

o Is the Bureau on the right trajectory to be able to implement these plans?

No. In February 2015, we reported that key challenges needed to be addressed to successfully enable Internet response for the 2020 Census—a critical component to generating the largest possible self-response.⁸ Specifically, the Bureau had not yet determined how key IT research questions related to estimating the Internet self-response rate and determining the IT infrastructure needed to support the large volume of data processing and storage expected for the 2020 Census were to be answered. We recommended that the Bureau, among other things, develop methodologies and plans for answering these key questions in time to inform upcoming design decisions. The Bureau subsequently released 2020 Census design decisions in October 2015 that deferred answering these questions, as well as other Internet response-related decisions, such as whether the Bureau will provide a mobile application and what languages would be available, until 2016 through 2017.

Also in our February report, we stated that high-level time frames had not been established for making decisions related to implementing cloud computing, which was intended to help the Bureau scale up its IT infrastructure for the 2020 Census, and we recommended that the Bureau develop these time frames. To the Bureau's credit, the 2020 Census Operational Plan established these high-level cloud testing and readiness milestones. However, these milestones include completing an alternatives assessment of the solutions architecture and acquiring cloud computing services for each of the Census tests from 2016 through 2018—which as previously mentioned leaves a narrow window for ensuring that all production systems are in place for end-to-end testing.

⁷GAO, *Information Security: Actions Needed by Census Bureau to Address Weaknesses*, GAO-13-63 (Washington, D.C.: Jan. 22, 2013). Another version of this report was issued for limited distribution.

⁸GAO, *2020 Census: Key Challenges Need to Be Addressed to Successfully Enable Internet Response*, GAO-15-225 (Washington, D.C.: Feb. 5, 2015).

○ **Are the right deadlines in place for decision-making?**

No. We are concerned about the deferred timing of many of the remaining decisions and their impact on delivering IT solutions for 2020, including the Internet self-response option, for the reasons discussed in my prior responses. Accordingly, the Bureau should take swift action to address the open recommendations in our February report in order to more effectively carry out its redesign plans.

Bureau's Use of Administrative Records and Additional Resources

10. What are some key security issues with using administrative records, and what recommendations would you make in remedying that?

We have previously reported that until the Bureau fully implements all of our security recommendations from our January 2013 report,⁹ it will have limited assurance that its information and systems are being adequately protected against unauthorized access, use, disclosure, modification, disruption, or loss.¹⁰ Specifically, as previously mentioned, we made 115 recommendations aimed at addressing IT security weaknesses—66 of which have been fully implemented as of October 2015.

In addition, our previous reports, and those by federal inspectors general, describe persistent information security weaknesses that place federal agencies, including the Bureau, at risk of disruption, fraud, or inappropriate disclosure of sensitive information. In particular, in January 2014, we reported about the risk to individual privacy associated with data sharing across government agencies, including individuals' personally identifiable information, such as names and Social Security numbers.¹¹ One important analytical technique used by agencies to enhance their ability to share data and evaluate program performance, among other things, is computer matching—a term commonly used to refer to the computerized comparison of information, generally including personally identifiable information, in two or more information systems. While computer matching programs have been successful in identifying fraud, waste, and abuse in federal benefit programs, we reported that if proper controls are not in place, they can also adversely affect the privacy and due process rights of individuals whose records are being matched. We have an ongoing review of the CEDCAP program that will allow us to better understand the specific IT security issues related to the 2020 Census design, including those related to the use of administrative records. Our tentative plan is to issue a report in spring 2016, potentially including recommendations if weaknesses are found.

11. Did the Bureau's lack of addressing key IT issues affect the secure use of administrative records in its 2015 tests?

We have not conducted the necessary work to answer this question.

⁹GAO-13-63. Another version of this report was issued for limited distribution.

¹⁰GAO, *2020 Census: Additional Actions Would Help the Bureau Realize Potential Administrative Records Cost Savings*, GAO-16-48 (Washington, D.C.: Oct. 20, 2015).

¹¹GAO, *Computer Matching Act, OMB and Selected Agencies Need to Ensure Consistent Implementation*, GAO-14-44 (Washington, D.C.: Jan. 13, 2014).

12. In preparation for the 2016 and 2017 tests, does GAO recommend the Bureau to use administrative records differently than it did in its 2015 tests? If so, how?

Please refer to the response of my colleague, Robert Goldenkoff, on this question. In addition, from an IT perspective, the Bureau should consider the findings contained in our January 2014 report cited in my response to question 10 to help inform test preparations.

13. The Bureau seeks to use additional sources it currently does not have access to, such as NDNH and KidLink.

○ **Do you recommend the Bureau uses them? Why or why not?**

Please refer to the response of my colleague, Robert Goldenkoff, on this question. Additionally, from an IT perspective, in light of the scope and complexity of the planned 2020 IT systems to be implemented in a narrowing window of time available, the Bureau should consult with the Commerce Chief Information Officer to fully understand the IT implications associated with acquiring these additional databases (e.g., information security measures and adapting the 2020 IT architecture and supporting systems to integrate the use of these additional data sources), and then make an informed decision whether to continue or end pursuit of these records.

○ **What potential security issues will this raise?**

Please refer to my responses to questions 8 and 10.